

Mar 25 02:47

US-08-644-289-3.rst

1

WIRELESS

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Merch_mn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Mar 25 02:48:13 1997; Maspar time 60.19 Seconds

Tabular output not generated.

Title: >US-08-644-289-3
Description: (1-28) from US08644289.seq
Perfect Score: 28
N.A. Sequence: 1 AGTCGAATTCATTCGACCAATCCCTGCCT 28
Comp: TCAGCTTAAGTAACCTCGTAGACCGCA

Scoring table: TABLE default
Gap 10

Mmatch STD : Dbase 0; Query 0

Searched: 630489 seqs, 220513910 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99
EST-STS-TMO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129

Mar 25 02:47

US-08-644-289-3.rst

2

130:STS2 131:STS3 132:STS4 133:STS5 134:STS6 135:STS7
136:STS8 137:STS9 138:STS10 139:STS11 140:STS12
141:STS13 142:STS14 143:STS15 144:STS16 145:STS17
146:STS18 147:STS19 148:STS20 149:STS21 150:STS22
151:STS23 152:STS24 153:STS25 154:STS26 155:STS27
156:STS28 157:STS29 158:STS30 159:STS31 160:STS32
161:STS33 162:STS34 163:STS35 164:STS36 165:STS37
166:STS38 167:STS39 168:STS40 169:STS41 170:STS42
171:STS43 172:STS44 173:STS45 174:STS46

Statistics: Mean 7.030; Variance 1.506; scale 4.668

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	ID	Description	Pred. No.
1	19	67.9	410	62	N51985	y207f07.sl Homo sapie	4.01e-06
2	18	64.3	430	115	T52639	ya72a11.sl Homo sapie	7.88e-05
3	18	64.3	457	98	R89652	ym7401.sl Homo sapie	7.88e-05
4	18	64.3	473	34	H90484	yv01g06.sl Homo sapie	7.88e-05
5	18	64.3	509	17	H38514	yp69b12.sl Homo sapie	7.88e-05
6	17	60.7	197	128	T99302	ye63a10.sl Homo sapie	1.41e-03
7	17	60.7	341	105	T11785	Al138R Homo sapiens c	1.41e-03
8	17	60.7	386	72	R02746	ye76d09.sl Homo sapie	1.41e-03
9	17	60.7	443	157	HS020330	za71e11.sl Homo sapie	1.41e-03
10	17	60.7	443	67	N69020	zb03c03.sl Homo sapie	1.41e-03
11	16	57.1	202	164	HS681308	zb14f11.sl Homo sapie	2.27e-02
12	16	57.1	215	163	HS611312	H. sapiens partial cD	2.27e-02
13	16	57.1	225	39	HS0KIC122	hbc022 Homo sapiens c	2.27e-02
14	16	57.1	252	105	T10560	Human STS UT854, 5' p	2.27e-02
15	16	57.1	284	138	H0MPT854A	yu84e11.sl Homo sapie	2.27e-02
16	16	57.1	285	105	T11542	Human STS UT854, 5' p	2.27e-02
17	16	57.1	296	30	H78195	KK6769P Homo sapiens	2.27e-02
18	16	57.1	319	142	N84745	Human fetal brain cDN	2.27e-02
19	16	57.1	330	49	H0M418C11B	H. sapiens (DXS7105) D	2.27e-02
20	16	57.1	344	174	HS4152M05	yr20b08.sl Homo sapie	2.27e-02
21	16	57.1	347	24	HS8307	yb59f05.sl Homo sapie	2.27e-02
22	16	57.1	354	117	T59311	yp66e04.sl Homo sapie	2.27e-02
23	16	57.1	359	21	H50664	EST03150 Homo sapiens	2.27e-02
24	16	57.1	365	104	T05261	yt66d09.sl Homo sapie	2.27e-02
25	16	57.1	402	99	R93365	yu15b12.sl Homo sapie	2.27e-02
26	16	57.1	407	26	H65947	yu18c08.sl Homo sapie	2.27e-02
27	16	57.1	407	28	H70615	yz42d06.sl Homo sapie	2.27e-02
28	16	57.1	411	22	H53572	yn01e12.sl Homo sapie	2.27e-02
29	16	57.1	414	67	N67592	yt74e04.sl Homo sapie	2.27e-02
30	16	57.1	419	98	R90920	N18b10 Homo sapiens c	2.27e-02
31	16	57.1	424	36	H97697	za11f02.sl Homo sapie	2.27e-02
32	16	57.1	427	65	N59638	ym53c06.sl Homo sapie	2.27e-02
33	16	57.1	427	106	T17269	yz48d11.sl Homo sapie	2.27e-02
34	16	57.1	428	159	HS156315	ye55g06.sl Homo sapie	2.27e-02
35	16	57.1	428	67	N68188	yv45c05.sl Homo sapie	2.27e-02
36	16	57.1	432	55	N29941	ys52d12.sl Homo sapie	2.27e-02
37	16	57.1	433	33	H87153		
38	16	57.1	439	67	N66945		
39	16	57.1	447	128	T97624		
40	16	57.1	449	18	H39611		
41	16	57.1	494	70	N76766		
42	16	57.1	515	52	N20867		

Mar 25 02:47

US-08-644-289-3.rst

3

43 16 57.1 550 62 N50065 yz10h03.s1 Homo sapie 2.27e-02
c 44 16 57.1 578 34 H93071 yv06h04.s1 Homo sapie 2.27e-02
c 45 16 57.1 1953 151 W18192 IMAGE:20083 Soares In 2.27e-02

ALIGNMENTS

RESULT 1
LOCUS N51985 410 bp mRNA EST 14-FEB-1996
DEFINITION y207f07.s1 Homo sapiens cDNA clone 282373 3' similar to contains
Alu repetitive element; contains element LTR1 repetitive element ;.
ACCESSION N51985
NID g1193151
KEYWORDS EST.
SOURCE human clone=282373 primer=m13 -40 forward library=Soares multiple
sclerosis 2NBHSP vector=pT730 (Pharmacia) with a modified
polylinker V TYPE: phagemid host=DH10B (ampicillin resistant)
Rs1el=Not I Rs1el2=Eco RI 46 year old male. 1st strand cDNA was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCATCTGAGTGGAGCGCGCATTTTCTTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7 vector (Pharmacia). Library went
through one round of normalization to a Cot = 5. Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4
multiple sclerosis lesions from one patient was kindly provided by
Dr. Kevin G. Becker (NINDS/NIH).

ORGANISM

Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tectipoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 410)
Hiller, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

REFERENCE

AUTHORS

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 188
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
FEATURES
source
1..410
/organism="Homo sapiens"
/clone="282373"
/note="human"
<1..>410

mRNA
BASE COUNT 103 a 108 c 88 g 109 t 2 others
ORIGIN

Query Match 67.9%; Score 19; DB 62; Length 410;
Best Local Similarity 84.6%; Pred. No. 4.01e-06;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Mar 25 02:47

US-08-644-289-3.rst

4

Db 187 gccagatgctcattcaattgac 212
|||||
Cp 27 gccagatgctcattcaattgac 2

RESULT 2
LOCUS T52639 430 bp mRNA EST 08-FEB-1995
DEFINITION ya72a11.s1 Homo sapiens cDNA clone 67196 3' contains Alu repetitive
element.
ACCESSION T52639
NID g654499
KEYWORDS EST.
SOURCE human clone=67196 library=Stratagene placenta (#937225)
vector=pBluescript SK- host=501R cells (kanamycin resistant)
primers=21m13 Rs1el=EcoRI Rs1el2=XhoI Placental tissue from a
Caucasian male. Cloned unidirectionally. Primer: Oligo dT. Average
Insert size: 1.2 kb; Uni-ZAP XR Vector; 5' adaptor sequence:
5'-CGATTCGCGACGAC-3'; 3' adaptor sequence:
5'-CTCGAGTTTTTTTTTTTTTTT-3'.

ORGANISM

Homo sapiens

Eukaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 430)
Hiller, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevasakis, E.,
Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
WashU-Merck EST Project
Unpublished (1995)
COMMENT

TITLE Contact: Wilson RK
JOURNAL WashU-Merck EST Project
Unpublished (1995)
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 353
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
FEATURES
source
1..430
/organism="Homo sapiens"
/clone="67196"
/note="human"

BASE COUNT 91 a 119 c 104 g 112 t 4 others
ORIGIN
Query Match 64.3%; Score 18; DB 115; Length 430;
Best Local Similarity 100.0%; Pred. No. 7.88e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 186 agccagatgctccaat 203
|||||
Cp 28 acccagatgctccaat 11

RESULT 3
LOCUS R89652 457 bp mRNA EST 24-AUG-1995
DEFINITION ym91d01.r1 Homo sapiens cDNA clone 166849 5' similar to contains
Alu repetitive element; contains MER22 repetitive element ;.
ACCESSION R89652

NID 9954479
KEYWORDS EST.
SOURCE human clone=166849 library=Soares adult brain N2b4HB5Y

ORGANISM

REFERENCE AUTHORS

TITLE	COMMENT
The WashU-Merck EST Project Unpublished (1995)	

Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence stops: 303
 Source: IMAGE Consortium, LInL
 This clone is available royalty-free through LInL; contact the
 IMAGE Consortium (info@image.liml.gov) for further information.
 Location/Qualifiers

Query Match	64.3%;	Score 18;	DB 98;	Length 457;
Best Local Similarity	87.5%;	Pred. No. 7.88e-05;		
Matches	21; Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;

```

Db      172  gaatgcagctggaacacatcctgct 195
          |||| | ||| ||||| |||||
Qy      5   GAATTCATTGGGACATCTGCT 28

```

LOCUS	H90484	473 bp	mRNA	EST	28-NOV-1995
DEFINITION	YV01.g06.r1 Homo sapiens cDNA clone 241498 5' similar to confamilins				
ACCESSION	Alu repetitive element;.				
	H90484				

ORGANISMS

REFERENCE
AUTHORS

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 371
Source: IMAGE Consortium, LNL.
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.

Query Match	64.3%;	Score 18;	DB 34;	Length 473;
Best Local Similarity	87.5%;	Pred. No. 7.88e-05;		
Matches	21;	Conservative	0;	Mismatches 3;
			Indels	0;
			Gaps	0

Db	94	agccaggatggtgcacacgattc	117
Cp	28	AGCCAGATGCTCCCATGCATTTC	5
RESULT	5		
LOCUS	H38514	509 bp	mRNA
DEFINITION	yp65b12.r1 Homo sapiens cDNA clone 192671 5'		EST 16-AUG-1995
	Alu repetitive element; contains TIR repetitive element ;.		
ACCESSION	H38514		

Mar 25 02:47

US-08-644-289-3.Fst

7

Mar 25 02:47

US-08-644-289-3.rtf

88

NID	g908013
KEYWORDS	EST.
SOURCE	human clone=192671 library=Scares adult brain N2b4H855Y

Homo sapiens

Mar 25 02:47

US-08-644-289-3.rst

9

SOURCE human clone-A1138 library=Heart vector=Lambda gcl1 host=E.coli
Y1090 primer=GACACGACCACTGCTATG Ralte1=EcORI Ralte2=EcORI.

ORGANISM Homo sapiens
Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 341)
AUTHORS Liew,C.C., Hwang,D.M., Fung,Y.W., Laurensen,C., Cukerman,E.,
Tsui,S. and Lee,C.Y.
A catalogue of genes in the cardiovascular system as identified by
expressed sequence tags

TITLE Proc. Natl. Acad. Sci. USA 91, 10645-10649 (1994)
JOURNAL 95024171
MEDLINE
COMMENT Other ESTs: A1138F
Contact: Liew CC
Molecular Cardiology
University of Toronto
Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
Tel: 416/978/758
Fax: 416/978/5650
Email: liewcc@utcc.utoronto.ca.

FEATURES
source Location/Qualifiers
1..341
/organism="Homo sapiens"
/clone="A1138"
/note="human"

BASE COUNT 67 a 96 c 85 g 93 t
ORIGIN

Query Match 60.7%; Score 17; DB 105; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.41e-03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 194 gccagatgtcccaat 210
|||||
Cc 27 gccacgatgctcccaat 11

RESULT 8
LOCUS R02746 386 bp mRNA EST 31-MAR-1995
DEFINITION yel609.s1 Homo sapiens cDNA clone 123665 3'.
ACCESSION R02746
NID g752482
KEYWORDS EST.
SOURCE human clone=123665 library=Soares fetal liver spleen INFLS
vector=pVT73D (Pharmacia) with a modified polylinker host=DH10B
(ampicillin resistant) primer=21m3 Ralte1=Pac I Ralte2=Eco RI
Liver and spleen from a 20 week-post conception male fetus. 1st
strand cDNA was primed with a Pac I - oligo(dT) primer [5'
AAGTCGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'], double-stranded
cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
I and cloned into the Pac I and Eco RI sites of the modified pVT73
vector. Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldi.

ORGANISM Homo sapiens
Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 386)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)

Mar 25 02:47

US-08-644-289-3.rst

10

COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 316
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source Location/Qualifiers
1..386
/organism="Homo sapiens"
/clone="123665"
/note="human"

BASE COUNT 109 a 86 c 68 g 123 t
ORIGIN

Query Match 60.7%; Score 17; DB 72; Length 386;
Best Local Similarity 84.0%; Pred. No. 1.41e-03;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 338 gtccaatcatgggccatgctg 362
||| || ||||| ||||| |||||
Qy 2 gtccatctattggacacatctctgc 26

RESULT 9
ID HS020330 standard; RNA; EST; 443 BP.
AC N69020;
DT 14-MAR-1996 (Rel. 47, Created)
DT 14-MAR-1996 (Rel. 47, Last updated, Version 1)
DE za7te11.s1 Homo sapiens cDNA clone 298028 3' similar to contains
DE Alu repetitive element;.
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
RN [1]
RP 1-443
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
The WashU-Merck EST Project;
RT Unpublished.
RL Contact: Wilson RK WashU-Merck EST Project Washington University
School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
est@wustl.wustl.edu High quality sequence stops: 340 Source: IMAGE
Consortium, LNL This clone is available royalty-free through LNL
CC; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC information. NCBI gi: 1225181
FH Key Location/Qualifiers
FH
FT source 1..443
FT /organism="Homo sapiens"
FT /clone="298028"
FT /note="human"
FT mRNA <1..443
FT Sequence 443 BP; 100 A; 108 C; 89 G; 145 T; 1 other;

Mar 25 02:47

US-08-644-289-3.rst

11

Query Match 60.7%; Score 17; DB 157; Length 443;

Best Local Similarity 100.0%; Pred. No. 1,41e-03;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 328 gccagatggtcccaat 344

Cp 27 gccagatggtcccaat 11

RESULT 10

LOCUS N69020 443 bp mRNA EST 13-MAR-1996

DEFINITION zaf1el.s1 Homo sapiens cDNA clone 298028.3 similar to contains

Alu repetitive element;.

ACCESSION N69020

NID g1225181

KEYWORDS

SOURCE

human clone=298028 primer=mi3 -40 forward library=Soares fetal lung
NBHL19M vector=pT7T3D (Pharmacia) with a modified polylinker
host=DM10B (ampicillin resistant) Reiter-Not I Reiter-Eco RI 19
week fetus. 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5'-TGTACCAATCTGACAGCGAGCCGCGATTTTCTTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T3 vector (Pharmacia). Library went
through one round of normalization to a Cot = 5. Library
constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 443)

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)

TITLE

JOURNAL

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 340
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.jnl.gov) for further information.

FEATURES

source 1..443

/organism="Homo sapiens"

/clone="298028"

/note="human"

<1..>443

BASE COUNT

100 a 108 c 89 g 145 t 1 others

ORIGIN

mRNA

Query Match 60.7%; Score 17; DB 67; Length 443;

Best Local Similarity 100.0%; Pred. No. 1,41e-03;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 328 gccagatggtcccaat 344

Cp 27 gccagatggtcccaat 11

Mar 25 02:47

US-08-644-289-3.rst

12

RESULT 11

ID HS681308 standard; RNA; EST; 202 BP.

AC N80681;

DT 04-APR-1996 (Rel. 47, Created)

DT 05-APR-1996 (Rel. 47, Last updated, Version 2)

DE zb03c03.s1 Homo sapiens cDNA clone 300964.3.

KM EST.

OS Homo sapiens (human)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Primates; Haplorhina; Catarrhini; Hominoidea.

RN [1]

RP 1-202

RA Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

RA Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

RA Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,

RA Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., Wilson, R.,

RT "The WashU-Merck EST Project";

RL Unpublished.

Contact: Wilson RK WashU-Merck EST Project Washington University
School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:

est@wustl.wustl.edu High quality sequence stops: 163 Source: IMAGE
Consortium, LNL This clone is available royalty-free through LNL

; contact the IMAGE Consortium (info@image.jnl.gov) for further
information. NCBI gi: 1243382

CC

FH Key Location/Qualifiers

FH

FT source

1..202

/organism="Homo sapiens"

/clone="300964"

/note="human"

FT mRNA

FT

SQ Sequence 202 BP; 37 A; 82 C; 40 G; 42 T; 1 other;

Query Match 57.1%; Score 16; DB 164; Length 202;

Best Local Similarity 86.4%; Pred. No. 2.27e-02;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 152 ccagatggtcccaatgaagc 173

Cp 26 ccagatggtcccaatgaatgc 5

RESULT 12

ID HS611312 standard; RNA; EST; 215 BP.

AC N79611;

DT 04-APR-1996 (Rel. 47, Created)

DT 05-APR-1996 (Rel. 47, Last updated, Version 2)

DE zb14f11.s1 Homo sapiens cDNA clone 302061.3.

KM EST.

OS Homo sapiens (human)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Primates; Haplorhina; Catarrhini; Hominoidea.

RN [1]

RP 1-215

RA Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

RA Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

RA Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,

RA Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., Wilson, R.,

RT "The WashU-Merck EST Project";

RL Unpublished.

Contact: Wilson RK WashU-Merck EST Project Washington University

CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
 CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
 CC est@stlscn.wustl.edu High quality sequence stops: 193 Source: IMAGE
 CC Consortium, LNL. This clone is available royalty-free through LNL
 CC ; contact the IMAGE Consortium (info@image.lnl.gov) for further
 CC information. NCBI gi: 1242312

FH Key Location/Qualifiers

FT source 1..215

FT /organism="Homo sapiens"

FT /clone="302061"

FT /note="human"

FT mRNA <1..>215

SQ Sequence 215 BP; 42 A; 86 C; 41 G; 44 T; 2 other;

Query Match 57.1%; Score 16; DB 163; Length 215;

Best Local Similarity 86.4%; Pred. No. 2.27e-02;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 151 ccagcatggtccactgaagtc 172

Cp 26 CCAGCATGTCCTCATGATTC 5

RESULT 13

LOCUS HSC0K122 225 bp RNA EST 28-JAN-1995

DEFINITION H. sapiens partial cDNA sequence; clone c-0kcl2.

ACCESSION F01978

NID 9645535

KEYWORDS partial cDNA sequence; transcribed sequence fragment.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotes; mitochondrial eukaryotes; Metazoa/Eumycota group;
 Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
 Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
 Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
 Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 225)

Genexpress.

Direct Submission

Submitted (19-JAN-1995) to the EMBL/GenBank/DDBJ databases.

Genethon, B.P. 60, 91002 Evry Cedex France and Genetique
 Molculaire et Biologie du developpement, CNRS UPR420 B.P. 8, 94801
 Villejuif Cedex France. E-mail: genexpress@genethon.fr

REFERENCE 2 (bases 1 to 225)

Genexpress.

The Genexpress cDNA program

Unpublished

3 (bases 1 to 225)

Auffray, C., Behar, G., Boile, F., Boucher, C., da Silva, C.,
 Devignes, M.D., Duprat, S., Houllgate, R., Jumeau, M.N., Lamy, B.,
 Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pletu, G., Pouliot, Y.,
 Sebastiani-Kabakchie, C. and Tessier, A.

IMAGE: Integrated molecular analysis of the human genome and its
 expression

C. R. Acad. Sci., III, Sci. Vie 318, 263-272 (1995)

Cloning method: total mRNA was oligo-(dT) primed and directionally
 cloned 5' -> 3' into the HindIII -> NotI sites of the lambda BA
 vector;

Sequencing method: single read, full automatic;

Primer: (-21)M13 universal;

cDNA sequence complementary to mRNA (3'end)

Stretch removed: removed at sequence 5'end

Normalization method: Bento Soares, P.N.A.S in press;

Genexpress_library_id: C;
 Genexpress sequence id: alc-0kcl2.
 Location/Qualifiers

FEATURES

source

1..225

/organism="Homo sapiens"

/clone lib="normalized infant brain cDNA from B.Souares,
 Psychiatry Dept. Columbia University USA"

/sex="female"

/tissue type="total brain"

/dev stage="3 months old"

/isolate="muscular atrophy patient"

BASE COUNT 59 a 56 c 43 g 67 t

Query Match 57.1%; Score 16; DB 39; Length 225;

Best Local Similarity 94.4%; Pred. No. 2.27e-02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 34 agccagatggtctcaat 51

Cp 28 AGCCAGATGTCCTCAAT 11

RESULT 14

LOCUS T10560 252 bp mRNA EST 29-NOV-1993

DEFINITION hbc022 Homo sapiens cDNA clone hbc022 5'end.

ACCESSION T10560

NID g390714

KEYWORDS EST.

Human clone=hbc022 library=Human pancreatic islet vector=Lambda

ZAP11 primer=5K primer Rsite1=Eco RI Rsite2=Xho I mRNA was prepared

from normal adult human islets. cDNA was directionally synthesized

from the Xho I in the vector to the EcoRI site. cDNA was size

fractionated to remove sequences <1000 bp in size.

Homo sapiens

Eucaryotes; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;

Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 252)

Bell, G.I. and Takeda, J.

A molecular inventory of human pancreatic islets: sequence analysis
 of 1000 cDNA clones
 Hum. Mol. Genet. 2, 1793-1798 (1993)

COMMENT Contact: Bell GI or Takeda J

HMNI Univ. of Chicago

5841 S. Maryland Ave., MC1028, Chicago IL 60637

Tel: 312/7029116

Fax: 312/7020271

Email: g-bell@uchicago.edu.

Location/Qualifiers

FEATURES

source

1..252

/organism="Homo sapiens"

/clone="hbc022"

/note="human"

BASE COUNT 52 a 70 c 63 g 67 t

ORIGIN

Query Match 57.1%; Score 16; DB 105; Length 252;

Best Local Similarity 94.4%; Pred. No. 2.27e-02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 39 agccagatggtctcaat 56

|||||

Mar 25 02:47

US-08-644-289-3.rst

15

Cp 28 AGCAGCATGCTCCCAAT 11

RESULT 15

LOCUS HUMUT854A 284 bp DNA STS 28-DEC-1994

DEFINITION Human STS UT854, 5' primer bind.

ACCESSION L31251

NTD 9605497

KEYWORDS PCR primer; STS sequence; microsatellite DNA; microsatellite marker; sequence tagged site; tetranucleotide repeat.

SOURCE Homo sapiens DNA.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.

AUTHORS 1 (bases 1 to 284)

Genetic and physical mapping of simple sequence repeat containing sequence tagged sites from the human genome

Unpublished (1994) See COMMENT for author address

Submitted by: Utah Center for Human Genome Research University of Utah, Dept. of Human Genetics

2160 Eccles Institute of Human Genetics

Salt Lake City, UT 84112

e-mail: str@corona.med.utah.edu

Primer A: AGCTCAGCATTGACAC

Primer B: ATTACAGTGTAGCCACTG

End to label: Primer B

PCR Profile:

Initial Denaturation: 94C 300sec

Cycle Denaturation Annealing Extension 5 94

C 10 sec. 62 C 10 sec. 72 C 20 sec. 31 94 C 10 sec.

58 C 10 sec. 72 C 20 sec. Mg++: 1.25 mM

Gel: Acrylamide 7%, Formamide 32%, Urea 34%

Alleles: 1.

FEATURES

source

1..284

/organism="Homo sapiens"

/sequenced_mol="DNA"

primer_bind

4..22

BASE COUNT 93 a 47 c 91 g 43 t 10 others

ORIGIN

Query Match

57.1%; Score 16; DB 138; Length 284;

Best Local Similarity 94.4%; Pred. No. 2.27e-02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 14 attgagacatctgct 31

QY 11 ATTGGAGCATCTCGCT 28

Search completed: Tue Mar 25 02:49:22 1997
Job time : 69 secs.

(11)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_n n.a. ~ n.a. database search, using Smith-Waterman algorithm

```
Run on: Tue Mar 25 02:42:30 1997; MasPar time 10.59 Seconds
```

Tabular output not generated.

Title: >US-08-644-289-2
Description: (1-38) from US08644289.seq

N.A. Sequence: 1 AGTCAGCCTTAGAGTTAAAGGATGCCCATGCTACAGA 38

Comp: TCAGTCCGGAATCTCAATTTCTACGGGTACGATGCTC

Scoring table: **TABLE default**

Gap 10

```
Nmatch  STD :  Dbase 0;  Query 0
```

Searched: 113505 seqs, 43611913 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22

Statistics: Mean 5.939; Variance 3.336; scale 1.780

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No
1	38	100.0	38	21	T10486	Human wild-type p53 g	2.73e-13
2	34	89.5	1173	12	D67883	Mouse p53 DNA.	1.1e-10
3	26	68.4	91	9	O51746	Oligonucleotide probe	1.47e-05
c	4	63.2	91	9	O51746	Oligonucleotide probe	2.46e-04
5	19	50.0	471	11	Q62363	Fragment coding for h	2.07e-01
6	19	50.0	984	11	Q62361	Fragment coding for h	2.07e-01
7	19	50.0	1062	11	Q62359	Fragment coding for h	2.07e-01
8	19	50.0	1181	15	Q97854	Human p53 cDNA.	2.07e-01

9	19	50.0	1182.2	T29719	Wild type p53 gene se	2.07e-01
	19	50.0	1182.3	Q22995	Sequence encoding 53	2.07e-01
10	19	50.0	1182.12	Q67884	Human p53 DNA.	2.07e-01
11	19	50.0	1483.12	Q67864	Vaccinia H6 promoter/	2.07e-01
12	19	50.0	1512.12	Q67863	Vaccinia H6 promoter/	2.07e-01
13	19	50.0	1740.4	Q27944	Mutant p53 polyprotein	2.07e-01
14	19	50.0	1740.4	Q27944	Mixed oligonucleotide	7.47e-01
15	18	47.4	39.7	Q51787	Base substituted E.co	2.62e+00
16	17	44.7	204.1	N81164	MoMuLV gag matrix p	8.93e+00
17	16	42.1	393.15	Q92899	MoMuLV mutated gag	8.93e+00
18	16	42.1	393.15	Q92900	MoMuLV mutated gag	8.93e+00
19	16	42.1	393.15	Q92900	MoMuLV gag/pol gene	8.93e+00
20	16	42.1	420.20	T07375	MoMuLV gag/pol gene	8.93e+00
21	16	42.1	1047.2	Q10572	Human Natriuretic Pe	8.93e+00
22	16	42.1	5292.19	T13392	Hybrid vector pSF2.	8.93e+00
23	16	42.1	5294.19	T13391	Hybrid vector pSF2.	8.93e+00
24	16	42.1	5363.19	T13390	Hybrid vector pSF1.	8.93e+00
25	16	42.1	5364.19	T13393	Hybrid vector pPM1.	8.93e+00
26	16	42.1	6365.7	Q41173	Plasmid IXSN-RI-112	8.93e+00
27	16	42.1	6505.19	T13394	Hybrid vector pSF-MDR	8.93e+00
28	16	42.1	8045.17	T12237	Retroviral vector MFG	8.93e+00
29	16	42.1	8202.18	T09280	Novel AMP/MCF virus	8.93e+00
30	16	42.1	8316.12	Q74450	myoD retroviral expe	8.93e+00
31	16	42.1	8332.20	T07374	MoMuLV gag/pol gene	8.93e+00
32	16	42.1	8332.20	T08452	MoMuLV gag/pol gene	8.93e+00
33	16	42.1	8388.13	Q78191	myoD retroviral expe	8.93e+00
34	16	42.1	9115.12	Q74449	myoD retroviral expe	8.93e+00
35	16	42.1	9629.18	T14600	pKClX-10M-CGF expres	8.93e+00
36	15	39.5	606.21	T11593	Bacillus sp. oligonuc	2.94e+01
37	15	39.5	1415.2	N71089	Sequence of human gra	2.94e+01
38	15	39.5	1521.2	N60937	Plasmid pBRZ insert	2.94e+01
39	15	39.5	1525.1	N81478	Sequence encoding hum	2.94e+01
40	15	39.5	1531.2	N60936	Plasmid pBR4 insert	2.94e+01
41	15	39.5	2634.6	Q39093	Streptomyces nodosus	2.94e+01
42	15	39.5	3070.2	N71090	Sequence of human gra	2.94e+01
43	15	39.5	3760.17	T09341	Tobacco mosaic virus	2.94e+01
44	15	39.5	4765.18	T08132	Protease gene.	2.94e+01
45	15	39.5	4922.21	T18877	Human TCL-1 gene.	2.94e+01

ALIGNMENTS

RESULT 1

ID	standard; DNA; 38 B
T10486	

DT 03-OCT-1996 (first entry)

Human wild-type p53 gene N-terminal sense RT-PCR primer.

proliferative disorder: neoplasia: diagnosis: cell|cwc|e:
 tumour suppressor; p53; alternative splicing; antibody; cancer;
 KW

reverse transcriptase polymerase chain reaction; ss.

05 Homo sapiens

PN CA2150994-A.

PD 15-DEC-1995.

PE 05-JUN-1995; 150994.
 DB 14-JUN-1994; 115-259613

PA (HEAT-) HEALTH RES INC
14-00N-1774, OS-207612
FN

PI Kulesz-Martin MF;

DR WPI; 96-140137/15.

PT New antibodies spe

PT useful in diagnosis

PS Dieckmann: Page 7

CC T10486-T10488 are

CC N-terminal (nucleo

CC gene. An epitope

CC (Q51135). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also Q51735-45 and Q51747-59.

SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 63.2%; Score 24; DB 9; Length 91;
Best Local Similarity 5.9%; Pred. No. 2,46e-04;
Matches 2; Conservative 27; Mismatches 5; Indels 0; Gaps 0;

Db 29 hhvvhvhhvvvhhvhhvhhvhhvhhvsvct 62

38 TCTGACGATGGGATCCTTACTAGGACCT 5

RESULT 5

ID Q62363 standard; cDNA; 471 BP.

AC Q62363;

DT 18-NOV-1994 (first entry)

DE Fragment coding for human p53 amino acids 237-393.

KM Human nuclear phosphoprotein p53; tumour suppressor gene product;

KW anti-oncogene; cancer; tumour; antibody binding region; epitope; ds.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..471

FT /tag= a

FT /product= p53(237-393)

FT /transl_except= pos:109..111,aa:Arg

FT /note= "partial CDS, therefore sequence does not

FT contain start or stop codons"

PN M09408241-A.

PD 14-APR-1994.

PF 30-SEP-1993; E02666.

PR 30-SEP-1992; DE-232823.

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM STIFTUNG.

PI Klein R, Schranz P, Tessmer C, Volkmann M, Zentgraf H;

DR WPI; 94-135732/16.

P-PSDB; R51878.

PT Non-radioactive detection of p53 specific antibodies - by capture

PT on immobilised p53 or its fragments, then reaction with labelled

PT second antibody, for diagnosis of tumours and suitable for

PT screening

PS Claim 13; Page 25; 35pp; German.

CC Antibodies specific for p53 are detected by binding to immobilised

CC fragments of the p53 gene product containing the antibody-binding

CC region. Preferred fragments contain amino acids 1-241, 40-349,

CC 40-393, 66-241, 66-393, 237-349, 237-393 and esp. 9-33, 37-52 or

CC 368-386. See Q62357-Q62366 for nucleic acid sequences coding for

CC each of these fragments.

SQ Sequence 471 BP; 128 A; 133 C; 132 G; 78 T;

Query Match 50.0%; Score 19; DB 11; Length 471;
Best Local Similarity 82.8%; Pred. No. 2,07e-01;

Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 329 agccttggaactcaagatgccagct 357

QY 5 AGCCTTAGAGTTAAAGATGCCATGCT 33

RESULT 6

ID Q62361 standard; cDNA; 984 BP.

AC Q62361;

DT 18-NOV-1994 (first entry)

DE Fragment coding for human p53 amino acids 66-393.

KM Human nuclear phosphoprotein p53; tumour suppressor gene product;

KW anti-oncogene; cancer; tumour; antibody binding region; epitope; ds.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..984

FT /tag= a

FT /product= p53(66-393)

FT /transl_except= pos:622..624,aa:Arg

FT /note= "partial CDS, therefore sequence does not

FT contain start or stop codons"

PN M09408241-A.

PD 14-APR-1994.

PF 30-SEP-1993; E02666.

PR 30-SEP-1992; DE-232823.

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM STIFTUNG.

PI Klein R, Schranz P, Tessmer C, Volkmann M, Zentgraf H;

DR WPI; 94-135732/16.

P-PSDB; R51876.

PT Non-radioactive detection of p53 specific antibodies - by capture

PT on immobilised p53 or its fragments, then reaction with labelled

PT second antibody, for diagnosis of tumours and suitable for

PT screening

PS Claim 13; Page 24; 35pp; German.

CC Antibodies specific for p53 are detected by binding to immobilised

CC fragments of the p53 gene product containing the antibody-binding

CC region. Preferred fragments contain amino acids 1-241, 40-349,

CC 40-393, 66-241, 66-393, 237-349, 237-393 and esp. 9-33, 37-52 or

CC 368-386. See Q62357-Q62366 for nucleic acid sequences coding for

CC each of these fragments.

SQ Sequence 984 BP; 224 A; 311 C; 259 G; 190 T;

Query Match 50.0%; Score 19; DB 11; Length 984;
Best Local Similarity 82.8%; Pred. No. 2,07e-01;

Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 842 agccttggaactcaagatgccagct 870

QY 5 AGCCTTAGAGTTAAAGATGCCATGCT 33

RESULT 7

ID Q62359 standard; cDNA; 1062 BP.

AC Q62359;

DT 18-NOV-1994 (first entry)

DE Fragment coding for human p53 amino acids 40-393.

KM Human nuclear phosphoprotein p53; tumour suppressor gene product;

KW anti-oncogene; cancer; tumour; antibody binding region; epitope; ds.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..1062

FT /tag= a

FT /product= p53(40-393)

FT /transl_except= pos:700..702,aa:Arg

FT /note= "partial CDS, therefore sequence does not

FT contain start or stop codons"

PN M09408241-A.

PD 14-APR-1994.

PF 30-SEP-1993; E02666.

PR 30-SEP-1992; DE-232823.

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM STIFTUNG.

PI Klein R, Schranz P, Tessmer C, Volkmann M, Zentgraf H;

DR WPI; 94-135732/16.

P-PSDB; R51874.

PT Non-radioactive detection of p53 specific antibodies - by capture
PT on immobilised p53 or its fragments, then reaction with labelled
PT second antibody, for diagnosis of tumours and suitable for
PT screening
PS Claim 13; Page 22; 35pp; German.
CC Antibodies specific for p53 are detected by binding to immobilised
CC fragments of the p53 gene product containing the antibody-binding
CC region. Preferred fragments contain amino acids 1-241, 40-349,
CC 40-393, 66-241, 66-393, 237-349, 237-393 and esp. 9-33, 37-52 or
CC 368-386. See 062357-Q62366 for nucleic acid sequences coding for
CC each of these fragments.
SQ Sequence 1062 BP; 246 A; 329 C; 279 G; 208 T;

Query Match 50.0%; Score 19; DB 11; Length 1062;
Best Local Similarity 82.8%; Pred. No. 2,07e-01;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 920 aggccttgaactcaagatgccagct 948
QY 5 AGGCCTTAGAGTTAAAGATGCCATGCT 33

RESULT 8

ID Q97854 standard; cDNA; 1181 BP.
AC Q97854;
DT 06-DEC-1995 (first entry)
DE Human p53 cDNA.
KW Ubiquitin-conjugating enzyme; p53 protein; cell cycle;
KW cell proliferation; cancer; psoriasis; fibrosis; de.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..1181
FT /tag= a
PN M09518974-A.
PD 13-JUL-1995.
PF 04-JAN-1995; U00164.
PR 04-JAN-1994; US-176937.
PR 23-MAY-1994; US-247904.
PR 27-MAY-1994; US-250795.
PR 13-SEP-1994; US-305520.
PA (MITO-) MITOTIX INC.
PI Cotarel G, Draetta G, Eckstein JW, Gyuris J, Rolfe M;
DR WPI; 95-255137/33.
DR P-PSDB; R19658.
PT Identifying inhibitors of ubiquitin mediated proteolysis of cell cycle
PT regulatory proteins - also new ubiquitin conjugating enzymes, their
PT related nucleic acid, vectors, antibodies etc., useful for regulating
PT e.g. cell proliferation
PS Disclosure; Page 105-106; 157pp; English.
CC Human p53 cDNA (given in Q97854) was amplified from a HeLa cell
CC cDNA library using the primers given in Q97852-53. The gene
CC was subcloned into a baculovirus vector for expression of
CC recombinant p53 in Sf9 insect cells for use as a component of
CC an in vitro ubiquitin conjugating system.
SQ Sequence 1181 BP; 275 A; 366 C; 306 G; 234 T;

Query Match 50.0%; Score 19; DB 15; Length 1181;
Best Local Similarity 82.8%; Pred. No. 2,07e-01;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1037 aggccttgaactcaagatgccagct 1065
QY 5 AGGCCTTAGAGTTAAAGATGCCATGCT 33

RESULT 9
ID T29719 standard; cDNA; 1182 BP.
AC T29719;
DT 29-OCT-1996 (first entry)
DE Wild type p53 gene sequence.
KW p53 gene; cancer; carcinoma; neoplastic; neoplasia; phenotype;
KW osteosarcoma cells; lung carcinoma cells; lymphoma cells;
KW leukemia cells; soft tissue sarcoma cells; breast cells;
KW bladder cells; prostate carcinoma cell; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..1182
FT /tag= a
FT /product= p53 protein.
FT misc difference 19..21
FT /tag= b
FT /transl_except= CAT encodes Aspartic acid.
PN EP-710722-A1.
PD 08-MAY-1996.
PF 23-AUG-1991; 307791.
PR 24-AUG-1990; US-573405.
PA (REGC) UNIV OF CALIFORNIA.
PI Chen P, Lee W;
DR WPI; 96-223439/23.
DR P-PSDB; R91933.

PT Use of wild-type p53 gene - in a medicament for suppressing the
PT neoplastic phenotype of a cancer cell lacking wild-type p53 protein
PS Claim 1; Page 5; 25pp; English.
CC The wild-type p53 gene can be used in the production of a medicament
CC for suppressing the neoplastic phenotype of a cancer cell lacking
CC endogenous wild type p53 protein. Cancer cells suppressed in such
CC fashion include osteosarcoma cells, lung carcinoma cells, lymphoma
CC cells, leukemia cells, soft tissue sarcoma cells or breast, bladder
CC or prostate carcinoma cells.
SQ Sequence 1182 BP; 278 A; 366 C; 304 G; 234 T;

Query Match 50.0%; Score 19; DB 22; Length 1182;
Best Local Similarity 82.8%; Pred. No. 2,07e-01;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1037 aggccttgaactcaagatgccagct 1065
QY 5 AGGCCTTAGAGTTAAAGATGCCATGCT 33

RESULT 10

ID Q22995 standard; DNA; 1182 BP.
AC Q22995;
DT 23-JUL-1992 (first entry)
DE Sequence encoding 53 KD cellular protein.
KW Cancer therapy; cancer suppressor gene; oncogenesis; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..1182
FT /tag= a
PN EP-475623-A.
PD 18-MAR-1992.
PF 23-AUG-1991; 307791.
PR 24-AUG-1990; US-573405.
PA (REGC) UNIV OF CALIFORNIA.
PI Lee WH, Chen P;
DR WPI; 92-090221/12.
DR P-PSDB; R22238.

PT Cloned p53 cDNA and protein prods. - for suppression of
 PT neoplastic phenotype e.g. In osteo-sarcoma(s), Leukaemia(s),
 PT lymphoma(s), etc.
 PS Claim 5; Page 15; 25pp; English.
 CC p53 cDNA, or its gene prods., can be used to suppress and eradicate
 CC cancers caused by defective, mutant or absent cancer suppressor
 CC genes. Variant forms of p53 are found in human breast, lung or
 CC colon carcinoma, lymphoma, leukaemia, etc., suggesting that mutation
 CC of the p53 genes is involved in oncogenesis. Specifically 213 Arg
 CC is replaced by 213 His, a mutation found exclusively in tumour cells.
 SQ Sequence 1182 BP; 277 A; 368 C; 303 G; 234 T;

Query Match 50.0%; Score 19; DB 3; Length 1182;
 Best Local Similarity 82.8%; Pred. No. 2.07e-01;
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Db 1037 aggccttggaactcaagatgcccaagct 1065
 QY 5 AGCCTTAGAGTTAAAGATGCCCATCTCT 33

RESULT 11
 ID 067884 standard; DNA; 1182 BP.
 AC 067884;
 DT 23-MAR-1995 (first entry)
 DE Human p53 DNA.
 KM Polymerase chain reaction; primer; amplify; NYVAC; ALVAC; recombinant;
 KM murine; interleukin-2; IL-2; pRM825; pmut-1; pBS-SK; pM151; TK vector;
 KM plasmid; vaccinia; H6 promoter; amplify; primer; antigenic response;
 KM polymerase chain reaction; poxvirus; pS0542; immunological response;
 KM pathogen; human; interferon; IFN; ss.
 OS Synthetic.
 PN W09416716-A.
 PD 04-AUG-1994.
 PF 21-JAN-1994; 000888.
 PR 21-JAN-1993; US-007115.
 PR 19-JAN-1994; US-184009.
 PA (VIR0-) VIRGENETICS CORP.
 PI Cox WI, Paolletti E, Tartaglia J;
 DR WPI; 94-263767/32.
 PT Attenuated recombinant virus used for cancer therapy - comprises
 PT DNA encoding cytokine and/or tumour associated antigen
 PS Example 32; Fig 39; 232pp; English.
 CC This sequence represents the wildtype human p53 gene from the translation
 CC initiation codon to the stop codon. This sequence was used in the
 CC construction of an ALVAC-based recombinant virus containing a mutant
 CC form of the human p53 gene. The mutant form has a G/A substitution at
 CC position 524, changing an Arg residue at position 175 to a His residue.
 CC The plasmid pM110 (see also 067864) contains the vaccinia H6 promoter
 CC and the wild type human p53 gene in the ALVAC C5 insertion site. The
 CC mutant p53 gene was obtained from plasmid Cx22A and cloned into pM110
 CC to generate pM143. Recombination between pM143 and ALVAC rescuing
 CC virus produced recombinant virus vCP270, which contains the vaccinia H6
 CC promoted mutated human p53 in the C5 locus. The resulting virus may be
 CC used in a composition for inducing an antigenic or immunological
 CC response, ie. for immunisation against pathogens.
 SQ Sequence 1182 BP; 276 A; 365 C; 307 G; 234 T;

Query Match 50.0%; Score 19; DB 12; Length 1182;
 Best Local Similarity 82.8%; Pred. No. 2.07e-01;
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Db 1037 aggccttggaactcaagatgcccaagct 1065
 QY 5 AGCCTTAGAGTTAAAGATGCCCATCTCT 33

QY 5 AGCCTTAGAGTTAAAGATGCCCATCTCT 33

RESULT 12
 ID 067864 standard; DNA; 1483 BP.
 AC 067864;
 DT 22-MAR-1995 (first entry)
 DE Vaccinia H6 promoter/human p53 expression cassette from vCP207.
 KM Polymerase chain reaction; primer; amplify; NYVAC; ALVAC; recombinant;
 KM human; p53; wild type; mutant; p53wtXba1SP6/T3; p53-21Xba1; p53-238Xba1;
 KM vaccinia; H6 promoter; plasmid; pRM825; expression cassette;
 KM ALVAC (COP) p53; poxvirus; antigenic response; immunological response;
 KM pathogen; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc feature 1..108
 FT /tag= a
 FT /note= "Flanking sequence"
 FT promoter 109..232
 FT /tag= b
 FT /note= "Vaccinia H6 promoter"
 FT CDS 233..1414
 FT /tag= c
 FT /product= Wild type human p53
 FT misc feature 1415..1483
 FT /tag= d
 FT /note= "Flanking sequence"
 PN W09416716-A.
 PD 04-AUG-1994.
 PF 21-JAN-1994; 000888.
 PR 21-JAN-1993; US-007115.
 PR 19-JAN-1994; US-184009.
 PA (VIR0-) VIRGENETICS CORP.
 PI Cox WI, Paolletti E, Tartaglia J;
 DR WPI; 94-263767/32.
 PT Attenuated recombinant virus used for cancer therapy - comprises
 PT DNA encoding cytokine and/or tumour associated antigen
 PS Example 15; Fig 18; 232pp; English.
 CC The sequences given in 067863-64 represent expression cassettes
 CC containing the vaccinia H6 promoter and the wild type human p53 gene
 CC in vP1101 and vCP207, respectively. These sequences were used in the
 CC construction of poxvirus based recombinant viruses expressing wild type
 CC and mutant forms of the human p53 gene product. The wild type and
 CC mutant genes were derived from plasmids p53wtXba1SP6/T3, p53-21Xba1 and
 CC p53-238Xba1 respectively. The vaccinia H6 promoter was derived from
 CC plasmid pRM825 and precisely linked to the 5'-most region of the p53
 CC gene. The expression cassettes generated by linking the vaccinia H6
 CC promoter and the p53 genes may be cloned in to the poxvirus insertion
 CC plasmid, ALVAC, to give vCP207, vCP193 and vCP191 recombinant
 CC viruses, and into NYVAC to give vP1101, vP1096 and vP1098. These
 CC viruses may be used in a composition for inducing an antigenic or
 CC immunological response, ie. for immunisation against pathogens.
 SQ Sequence 1483 BP; 367 A; 415 C; 372 G; 339 T;

Query Match 50.0%; Score 19; DB 12; Length 1483;
 Best Local Similarity 82.8%; Pred. No. 2.07e-01;
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Db 1269 aggccttggaactcaagatgcccaagct 1297
 QY 5 AGCCTTAGAGTTAAAGATGCCCATCTCT 33

RESULT 13

ID Q67863 standard; DNA; 1512 BP.
AC Q67863;
DT 22-MAR-1995 (first entry)
DE Vaccinia H6 promoter/human p53 expression cassette from vp1101.
KW Polymerase chain reaction; primer; amplify; NVVAC; ALVAC; recombinant;
human; p53; wild type; mutant; p53wtXbaISP6/T3; p53-21XbaI; p53-238XbaI;
KW vaccinia; H6 promoter; plasmid; pRM825; expression cassette;
KW ALVAC (CPP) p53; poxvirus; antigenic response; immunological response;
KW pathogen; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..144
FT /tag= a
FT /note= "Flanking sequence"
FT promoter 145..268
FT /tag= b
FT /note= "Vaccinia H6 promoter"
FT CDS 269..1450
FT /tag= c
FT /product= Human wildtype p53
FT misc_feature 1451
FT /tag= d
FT /note= "Flanking sequence"
PN M09416716-A.
PD 04-AUG-1994.
PF 21-JUN-1994; U00888.
PR 21-JUN-1993; US-007115.
PR 19-JAN-1994; US-184009.
PA (VIR0-) VIROGENETICS CORP.
PI Cox WI, Paolletti E, Tartaglia J;
DR WPI; 94-263767/32.
PT Attenuated recombinant virus used for cancer therapy - comprises
PT DNA encoding cytokine and/or tumor associated antigen
PS Example 15; Fig 17; 232pp; English.
CC The sequences given in Q67863-64 represent expression cassettes
CC containing the vaccinia H6 promoter and the wild type human p53 gene
CC in vp1101 and vCP207, respectively. These sequences were used in the
CC construction of poxvirus based recombinant viruses expressing wild type
CC and mutant forms of the human p53 gene product. The wild type and
CC mutant genes were derived from plasmids p53wtXbaISP6/T3, p53-21XbaI and
CC p53-238XbaI respectively. The vaccinia H6 promoter was derived from
CC plasmid pRM825 and precisely linked to the 5'-most region of the p53
CC gene. The expression cassettes generated by linking the vaccinia H6
CC promoter and the p53 genes may be cloned in to the poxvirus insertion
CC plasmid, ALVAC, to give vCP207, vCP193 and vCP191 recombinant
CC viruses, and into NVVAC to give vp1101, vp1096 and vp1098. These
CC viruses may be used in a composition for inducing an antigenic or
CC immunological response, ie. for immunisation against pathogens.
SQ Sequence 1512 BP; 379 A; 419 C; 380 G; 334 T;

Query Match 50.0%; Score 19; DB 12; Length 1512;
Best Local Similarity 82.8%; Pred. No. 2.07e-01;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

DB 1305 aggccttggaactcaagatgccagcgt 1333
||||| | | ||||| | | | | |
QY 5 AGCCTTAGAGTTAAAGATGCCATGCT 33







RESULT 14
ID Q27944 standard; cDNA; 1740 BP.
AC Q27944;
DT 11-FEB-1993 (first entry)
DE Mutant p53 polypeptide (hnp53HIS273).

KW Oncogene; tumour; neoplasia; ss.
OS Homo sapiens.
PN M09213970-A.
PD 20-AUG-1992.
PF 31-JAN-1992; U00878.
PR 01-FEB-1991; US-649566.
PR 21-JUN-1991; US-719172.
PA (ONCO-) ONCOGENE SCI INC.
PI Reynolds FH, Sorvillo JM, Stephenson JR, Zehab R;
DR WPI; 92-300054/36.
PT Diagnosing a neoplastic condition in a subject - by detecting the
PT presence of mutant p53 polypeptide in a fluid sample from the
PT subject
PS Disclosure; Page 52; 81pp; English.
CC Mutant p53 polypeptides encoded by activated p53 oncogenes are
CC present intracellularly at high concs. in mammalian tumours and
CC tumour cell lines. The correlation between the measured mutant p53
CC polypeptide conc. and the presence of neoplastic disease provides
CC for the early detection and monitoring of neoplastic conditions.
CC In the Experimental Detail section of the specification the cloning
CC and expression of mutant p53 polypeptide is described, where a cDNA
CC library was constructed using mRNA extracted from A431 cells (express
CC hnp53HIS273) as described by Harlow E. et al., (1985) J. Virol. 37:1601.
CC The sequence derived from activated p53 oncogene given in
CC the specification has been published by Harlow E. et al. (1985)
CC Molecular and Cellular Biology 5:1601-1610.
SQ Sequence 1740 BP; 400 A; 510 C; 457 G; 373 T;

Query Match 50.0%; Score 19; DB 4; Length 1740;
Best Local Similarity 82.8%; Pred. No. 2.07e-01;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

DB 1325 aggccttggaactcaagatgccagcgt 1353
||||| | | ||||| | | | | |
QY 5 AGCCTTAGAGTTAAAGATGCCATGCT 33

RESULT 15
ID Q51787 standard; DNA; 39 BP.
AC Q51787;
DT 20-DEC-1993 (first entry)
DE Mixed oligonucleotide #19 encodes ballast constituent.
KW Fusion protein; ballast constituent; pro-insulin production;
KW recombinant protein production; HMG CoA reductase;
KW human 3-hydroxy-3-methylglutaryl-coenzyme A-reductase;
KW mixed oligonucleotide; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT repeat_unit 4..6
FT /tag= a
FT /rpt_type= tandem
FT /note= "can be repeated y times, where y is 4-11,
FT provided that y+z is 6-12"
FT repeat_region 4..36
FT /tag= b
FT /note= "(DCD)11"
FT repeat_unit 36..38
FT /tag= c
FT /rpt_type= tandem
FT /note= "can be repeated z times, where z is 1-4,
FT provided that y+z is 6-12; N stands for
FT identical or different nucleotides,
FT excluding stop codons"
PN US5227293-A.

(a)  (b)  (c)  (d)  (e)  (f) 

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

```
MPsrch_pp  protein - protein database search, using Smith-Waterman algorithm
Run on:      Mon Mar 24 11:17:33 1997;  MasPar time 1.84 Seconds
Tabular output not generated.  111.735 Million cell updates/sec
```

Title:	US-08-644-289-1
Description:	(1-20) from US08644289.dep
Perfect Score:	151
Sequence:	1 SLRPFKALVREKGRPSHSC 20

Scoring table:

Searched: 88003 seqs, 10295656 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq25
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18

Statistics: Mean 20.568; Variance 59.693; scale 0.345

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No
1	151	100.0	20	18	R92698	Human wild-type p53	8.31e+01
2	60	39.7	247	3	R13252	Murine Cytotoxic Cell	1.53e+01
3	60	39.7	619	5	R27651	Human calcium channel	1.53e+01
4	60	39.7	993	5	R27650	Human calcium channel	1.53e+01
5	60	39.7	1833	13	R73055	Rabbit skeletal calci	1.53e+01
6	60	39.7	1873	1	P95645	Rabbit skeletal muscle	1.53e+01
7	60	39.7	1967	6	R33547	Sequence of the alpha	1.53e+01
8	60	39.7	2138	14	R72607	Human neuronal calciu	1.53e+01
9	60	39.7	2161	14	R71002	Human neuronal calciu	1.53e+01
10	60	39.7	2161	14	R71001	Human neuronal calciu	1.53e+01
11	60	39.7	2161	6	R33545	Sequence of the alpha	1.53e+01
12	60	39.7	2163	14	R71003	Human neuronal calciu	1.53e+01

Mat 24 11-15

US-08-644-289-1.rag

13	57	31.7	349	4	R20784	Tulip allylacyl amida	3.21e+01
14	57	37.7	553	3	R13984	Partial HIV ribonuc	3.21e+01
15	57	31.64	1944	6	R94446	Hepatitis GB virus (H	3.21e+01
16	55	36.4	685	2	R11331	Human Interleuking hor	5.22e+01
17	54	35.8	237	4	R20509	Human proteinase 3.	6.64e+01
18	54	35.8	256	15	R85639	M17 preproR-3.	6.64e+01
19	54	35.8	256	9	R45401	Deduced sequence of h	6.64e+01
20	54	35.8	677	18	R95992	Human choline acetyl	6.64e+01
21	54	35.8	885	16	R92747	Sf8 virus gB glycopro	6.64e+01
22	54	35.8	2813	3	P60653	Sequence of von Wille	6.64e+01
23	54	35.8	2813	3	P60653	Sequence of human von	6.64e+01
24	53	35.1	99	13	R72089	Abi SH2 region.	8.43e+01
25	53	35.1	148	6	R29180	Astrovirus serotype A	8.43e+01
26	53	35.1	399	7	R33934	Morphogen mop2 full l	8.43e+01
27	53	35.1	1011	18	R99638	Peripheral nervous sy	8.43e+01
28	53	35.1	1693	10	R51264	HEV strain protein en	8.43e+01
29	53	35.1	1835	17	R92316	Peripheral nervous sy	8.43e+01
30	53	35.1	1977	18	R99641	Peripheral nervous sy	8.43e+01
31	53	35.1	1884	18	R96639	Peripheral nervous sy	8.43e+01
32	53	35.1	1889	18	R99640	Peripheral nervous sy	8.43e+01
33	53	35.1	1889	18	R92317	Peripheral nervous sy	8.43e+01
34	53	35.1	2019	13	R65984	Cardiac sodium channe	8.43e+01
35	53	35.1	2020	2	R06813	Cardiac sodium channe	8.43e+01
36	52	34.4	223	7	R37573	Partial human skeleta	1.07e+02
37	52	34.4	436	17	R98143	The DNA polymerase.	1.07e+02
38	52	34.4	436	17	R98143	The DNA polymerase.	1.07e+02
39	52	34.4	690	3	R15482	Protein encoded by OR	1.07e+02
40	51	33.8	118	12	R67373	Heat shock protein HS	1.35e+02
41	51	33.8	118	13	R74338	Helicobacter pylori h	1.35e+02
42	51	33.8	305	17	R89750	AFT-1 (Arabidopsis fo	1.35e+02
43	51	33.8	385	14	R76578	Coarctan omega-12 de	1.35e+02
44	51	33.8	567	18	R81531	Pseudomonas fluoresce	1.35e+02
45	51	33.8	1196	16	R75189	Osteoinductive retrov	1.35e+02

ALIGNMENTS

RESULT 1
ID R92698 strand; peptide: 20 AA.
AC R92698;
DT 03-OCT-1996 (first entry)
DE Human wild-type p53 gene intron 10 encoded epitope.
KM Tumor suppressor; p53; alternative splicing; antibody; cancer;
KM proliferative disorder; psoriasis; prognosis; diagnosis; cell cycle.
OS Homo sapiens.
PN CA2150994-A.
PN 15-DEC-1995.
PF 05-JUN-1995; 150994.
PR 14-JUN-1994; US-259612.
PA (HEAL-) HEALTH RES INC.
PI Kulesz-Martin MF;
PI WPI; 96-140137/15.
PT New antibodies specific for alternatively spliced mammalian p53 -
PT useful in diagnosis or prognosis of cancer, and for establishing
PT individual treatment regimens.
PS Claim 5; Page 25; 40pp; English.
CS R92698 is an epitope from intron 10 of the human p53 gene. Antibodies
CC directed against this sequence in an alternatively spliced p53 are
CC useful in diagnosis and prognosis of human cancer. Antibodies against
CC the epitope may be used in the study of p53 and alternatively spliced
CC p53 (p53as) and their expression, when the antibodies are injected
CC into cells they may cause cell cycle arrest. Vectors contg. p53as
CC cDNA may be used in gene therapy of cancers and other proliferative
CC disorders e.g., psoriasis.

SQ Sequence 20 AA;

Query Match 100.0%; Score 151; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 8.31e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 slrpfkalvrekghrpshec 20

Oy 1 slrpfkalvrekghrpshec 20

RESULT 2

ID R13252 standard; Protein; 247 AA.

AC R13252;

DT 11-OCT-1991 (first entry)

DE Murine Cytotoxic Cell Protease-1.

KM mouse; CCP-1 inhibitor; cytotoxic T-lymphocytes; ss.

OS Mus musculus.

FH Key Location/Qualifiers

FT Peptide 1..20

FT /label= signal peptide

FT Protein 21..247

FT /label= CCP-1

PN M09110685-A.

PD 25-JUL-1991.

PF 17-JAN-1991; U00340.

PR 19-JAN-1990; US-467880.

PA (SERA-) SERAGEN INC.

PI Bleackley RC, Lobe CG, Paetkau VH, James MN, Murphy M;

DR WPI, 91-237889/32.

DR N-PSDB; Q12862.

PT DNA vectors, and inhibitors of cytotoxic cell protease - for

PT treatment of auto-immune diseases e.g. pernicious anaemia,

PT rheumatoid arthritis, allo-graft rejection etc.

PS Claim 5; Fig 3; 62pp; English.

CC The CCP-1 coding sequence was isolated from the cytotoxic T-cell

CC line MTL 2.8.2 generated from CBA/J mice. The amino acid sequence

CC of CCP-1 was predicted from the cDNA sequence. The structure of the

CC protein was used to design peptides which competitively inhibit the

CC protease. See also Q12863-6 and R13254-R13262.

SQ Sequence 247 AA;

Query Match 39.7%; Score 60; DB 3; Length 247;

Best Local Similarity 38.9%; Pred. No. 1.53e+01;

Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db 32 rpyrnallslkdqpeaic 49

Oy 3 rpfkalvrekghrpshec 20

RESULT 3

ID R27651 standard; Protein; 619 AA.

AC R27651;

DT 03-MAR-1993 (first entry)

DE Human calcium channel 27980/13.

KM Plasmid pCA3.3; Ca-flux assay.

OS Homo sapiens.

PN EP-507170-A.

PD 07-OCT-1992.

PF 23-MAR-1992; 104970.

PR 04-APR-1991; DE-110785.

PA (FARB) BAYER AG.

PI Franz J, Rae F, Unterebeck A, Weingaertner B;

DR WPI; 92-333446/41.

DR N-PSDB; Q29271.

PT Cloned human neuronal calcium channel sub-types - useful in

PT calcium flux assays to screen for neurone-specific calcium

PT channel ligands

PS Claim 2; Page 83-85; 101pp; German.

CC Human neuroblastoma cell line, hippocampus, frontal and temporal

CC cortex and visual cortex cDNA banks were screened with a probe

CC containing carp skeletal muscle Ca-channel cDNA. The cDNA clone

CC pCA3.3 was sequenced. The 5' end of the clone begins directly after

CC Domain I (amino acid 337) and the 3' end is at amino acid 922. The

CC neuronal calcium channel protein can be used for screening for Ca

CC channel ligands (agonists or antagonists). See also Q29259-Q29275.

SQ Sequence 619 AA;

Query Match 39.7%; Score 60; DB 5; Length 619;

Best Local Similarity 42.1%; Pred. No. 1.53e+01;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 598 lrplrainrakqlkhvqc 616

Oy 2 lrpfkalvrekghrpshec 20

RESULT 4

ID R27650 standard; Protein; 993 AA.

AC R27650;

DT 03-MAR-1993 (first entry)

DE Human calcium channel 27980/12.

KM Plasmid pCA3; Ca-flux assay.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc difference 42

FT /note= "corresponds to nonsense codon TAG"

FT Misc difference 47

FT /note= "corresponds to nonsense codon TGA"

PN EP-507170-A.

PD 07-OCT-1992.

PF 23-MAR-1992; 104970.

PR 04-APR-1991; DE-110785.

PA (FARB) BAYER AG.

PI Franz J, Rae P, Unterebeck A, Weingaertner B;

DR WPI; 92-333446/41.

DR N-PSDB; Q29270.

PT Cloned human neuronal calcium channel sub-types - useful in

PT calcium flux assays to screen for neurone-specific calcium

PT channel ligands

PS Claim 2; Page 78-82; 101pp; German.

CC Human neuroblastoma cell line, hippocampus, frontal and temporal

CC cortex and visual cortex cDNA banks were screened with a probe

CC containing carp skeletal muscle Ca-channel cDNA. The cDNA clone

CC pCA3 was sequenced. The 5' end of the clone lies between

CC Domains II and III. The clone comprises the complete Domains III

CC and IV and part of the C-terminus coding sequence. The human

CC neuronal calcium channel protein can be used for screening for Ca

CC channel ligands (agonists or antagonists). See also Q29259-Q29275.

SQ Sequence 993 AA;

Query Match 39.7%; Score 60; DB 5; Length 993;

Best Local Similarity 42.1%; Pred. No. 1.53e+01;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 254 lrplrainrakqlkhvqc 272

Oy 2 lrpfkalvrekghrpshec 20

OY 2 LPPKALVREKGRPSHSC 20

RESULT 5
ID R73055 standard; Protein; 1873 AA.
AC R73055;
DT 02-NOV-1995 (first entry)
DE Rabbit skeletal calcium channel (alpha)1-subunit.
KW Calcium channel; (alpha)1 subunit.
OS Oryctolagus cuniculus
PM Key Location/Qualifiers
FT Region 950..1100
FT /label= IIF7 monoclonal antibody epitope
FT Misc_difference 1808
FT /label= Thr
FT /note= "In Tanabe et al"
FT Misc_difference 1815
FT /label= Ala
FT /note= "In Tanabe et al"
FT Misc_difference 1835
FT /label= Ala
FT /note= "In Tanabe et al"
FT Modified_site 79
FT /label= Potential N-glycosylation site
FT Modified_site 257
FT /label= see above
FT Modified_site 797
FT /label= see above
FT Modified_site 1464
FT /label= see above
FT Modified_site 1674
FT /label= see above
FT Modified_site 687
FT /label= potential cAMP-dependent phosphorylation
FT Modified_site 1502
FT /label= see above
FT Modified_site 1575
FT /label= see above
FT Modified_site 1757
FT /label= see above
FT Modified_site 1772
FT /label= see above
FT Modified_site 1854
FT /label= see above
FT Modified_site 1552
FT /label= see above
FT Region 52..70
FT /label= transmembrane region
FT Protein 89..108
FT /label= see above
FT Region 121..139
FT /label= see above
FT Region 161..179
FT /label= see above
FT Region 199..218
FT /label= see above
FT Region 310..334
FT /label= see above
FT Region 433..451
FT /label= see above
FT Region 467..486
FT /label= see above
FT Region 495..513
FT /label= see above

FT Region 524..542
FT /label= see above
FT Region 562..581
FT /label= see above
FT Region 637..661
FT /label= see above
FT Region 800..818
FT /label= see above
FT Region 835..854
FT /label= see above
FT Region 867..885
FT /label= see above
FT Region 893..912
FT /label= see above
FT Region 931..950
FT /label= see above
FT Region 1041..1065
FT /label= see above
FT Region 1119..1137
FT /label= see above
FT Region 1153..1172
FT /label= see above
FT Region 1181..1199
FT /label= see above
FT Region 1232..1250
FT /label= see above
FT Region 1270..1289
FT /label= see above
FT /label= see above
PN US5407820-A.
PD 18-APR-1995.
PF 04-APR-1988; 176899.
PR 04-APR-1988; US-176899.
PR 04-APR-1988; US-603751.
PR 13-JUL-1992; US-914231.
PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
PI Brenner R, Ellis SB, Harpold MM, Schwartz A, Williams ME;
DR WPI; 95-161088/21.
DR N-PSDB; Q87978.
PT DNA encoding alpha-1 and alpha-2 calcium channel sub-unit(s) -
PT for detecting calcium channel agonists/antagonists or
PT Lambert-Eaton syndrome.
PS Disclosure; Fig 1; 32pp; English.
CC The cDNA sequence Q87978 is consistent with an approx 6,500 nt DHP
CC receptor (alpha)1 mRNA. It is 94% identical to the cDNA sequence
CC encoding the DHP receptor reported by Tanabe et al., Nature, 328:313-
CC 318, 1987. Nt differences were identified at 33 posns, of which
CC three (see FT) also result in AA changes (see R73055 FT). The
CC deduced AA sequence yields a calculated Mr of 212,143, in contrast to
CC the observed Mr 155K-170K, previously reported by others using SDS
CC polyacrylamide gel electrophoresis. The AA sequence is 99.8%
CC identical to that described by Tanabe et al. It contains four
CC internal repeated sequence regions. It has been proposed that the
CC segments of the four internally repeated regions represent 24
CC transmembrane segments. A eukaryotic cell expressing a chimaeric
CC cDNA is claimed, in which the first cDNA is selected from the gp,
CC consisting of cDNA that encodes a protein that has the AA sequence
CC in R73055 but with the Tanabe et al substitutions indicated in FT.
SQ Sequence 1873 AA;

Query Match 39.7%; Score 60; DB 13; Length 1873;
Best Local Similarity 42.1%; Pred. No. 1.53e+01;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
Db 902 lrplratnrakglkhvqc 920

Oy 2 LRPFKALVREKGRHPSHSC 20

||||:| | | : |

RESULT 6
ID p95645 standard; protein; 1873 AA.
AC p95645;
DT 21-MAR-1990 (first entry)
DE Rabbit skeletal muscle alpha-1 sub-unit gene product.
KW Skeletal muscle
OS Sylvilagus (rabbit).
PN W080909834-A.
PD 19-OCT-1989.
PF 04-APR-1988; 001408.
PR 04-APR-1988; US-176899.
PA (SALK) Salk Inst for Biol Stud.
PI Ellis SB, Williams ME, Harpold MM, Schwartz A, Sartor J;
DR WPI; 89-324236/44.
DR N-PSDB; N91778.
PT New DNA encoding alpha-2 subunit of animal calcium channel - also new
PT protein product and eukaryotic cells for testing cpds. for calcium
PT agonist or antagonist activity
PS Disclosure; page 16-1 to 18-3; 68pp; English.
CC Also used to diagnose Lambert-Eaton syndrome by reacting test serum
CC with alpha-1 and alpha-2 subunits. labelled fragments can be used as
CC probes.
SQ Sequence 1873 AA;

Query Match 39.7%; Score 60; DB 1; Length 1873;
Best Local Similarity 42.1%; Pred. No. 1.53e+01;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 902 LRPLRAINRAKGLKHVQC 920
Oy 2 LRPFKALVREKGRHPSHSC 20
||||:| | | : |

RESULT 7
ID R33547 standard; Protein; 1967 AA.
AC R33547;
DT 30-JUN-1993 (first entry)
DE Sequence of the alpha 1C human calcium channel subunit.
KW Human calcium channel subunit; diagnosis; agonist; antagonist;
KW Lambert Eaton syndrome.
OS Homo sapiens.
PN W09304083-A.
PD 04-MAR-1993.
PF 14-AUG-1992; 006903.
PR 15-AUG-1991; US-745206.
PR 10-APR-1992; US-868354.
PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
PI Brenner R, Ellis SB, Feldman DH, Harpold MM, Mccue AF,
DR Williams ME;
DR WPI; 93-093936/11.
DR N-PSDB; Q37813.
PT DNA encoding specific human calcium channel sub-units - used for
PT identifying calcium channel agonists and antagonists and
PT diagnosing Lambert Eaton syndrome
PS Disclosure; Page 102-109; 150pp; English.
CC Numerous alpha 1C-specific DNA clones were isolated.
CC Characterisation of the sequence revealed the alpha 1C coding
CC sequence, the alpha 1C initiation of translation sequence, and an
CC alternatively spliced region of alpha 1C. Q37814 and Q37815 encode
CC two possible amino terminal ends of the alpha 1C protein. Q37816

CC encodes an alternative exon for the IV S3 transmembrane domain.
SQ Sequence 1967 AA;

Query Match 39.7%; Score 60; DB 6; Length 1967;
Best Local Similarity 42.1%; Pred. No. 1.53e+01;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 1027 LRPLRAINRAKGLKHVQC 1045
Oy 2 LRPFKALVREKGRHPSHSC 20
||||:| | | : |

RESULT 8
ID R72607 standard; Protein; 2138 AA.
AC R72607;
DT 01-DEC-1995 (first entry)
DE Human neuronal calcium channel subunit alpha 1c-2.
KW Calcium channel subunit; antagonist; agonist; diagnosis;
KW Lambert Eaton Syndrome.
OS Homo sapiens.
PN W09504822-A.
PD 16-FEB-1995.
PF 11-AUG-1994; 009230.
PR 11-AUG-1993; US-105536.
PR 05-NOV-1993; US-149097.
PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
PI Ellis SB, Gillespie A, Harpold MM, Mccue AF, Williams ME;
DR WPI; 95-090900/12.
DR N-PSDB; Q87834.
PT DNA encoding human calcium channel sub-unit(s) - used for
PT developing prods. for studying calcium channels, e.g. for
PT obtaining agonists and antagonists
PS Claim 34; Page 260-269; 285pp; English.
CC Numerous alpha 1c-specific cDNA clones were isolated in order to
CC characterise the alpha 1c coding sequence, the initiation of
CC translation and an alternatively spliced region. Q84655 sets
CC forth one alpha 1c coding sequence (alpha 1c-1) and R71003 sets
CC out its deduced AA sequence. Q87834 and R72607 set out another
CC splice variant, designated alpha 1c-2. Q84656 encodes an
CC alternative exon for the IV S3 transmembrane domain. Other
CC alpha 1c variants can be constructed by selecting alternative
CC amino terminal ends in place of the ends in Q84655 and Q87834
CC and/or inserting the alternative exon in the appropriate
CC location (see Q84655 FT). In addition, a nt. sequence (see Q84655
CC FT) can be deleted or inserted to produce an alternative alpha 1c
CC splice variant.
SQ Sequence 2138 AA;

Query Match 39.7%; Score 60; DB 14; Length 2138;
Best Local Similarity 42.1%; Pred. No. 1.53e+01;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 1003 LRPLRAINRAKGLKHVQC 1021
Oy 2 LRPFKALVREKGRHPSHSC 20
||||:| | | : |

RESULT 9
ID R71002 standard; Protein; 2161 AA.
AC R71002;
DT 30-NOV-1995 (first entry)
DE Human neuronal calcium channel subunit alpha 1D including alternative.
DE exon encoding the 156 transmembrane domain.
KW Calcium channel subunit; antagonist; agonist; diagnosis;

KM Lambert Eaton Syndrome.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc_difference 373..406
 FT /label= encoded by alternative exon
 PN M09504822-A.
 PD 16-FEB-1995.
 PF 11-AUG-1994; U09230.
 PR 11-AUG-1993; US-105536.
 PR 05-NOV-1993; US-149097.
 PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
 PI Ellis SB, Gillespie A, Harpold WM, Mccue AF, Williams ME;
 DR N-PSDB; 084654.
 DR N-PSDB; 084654.
 PT DNA encoding human calcium channel sub-unit(s) - used for
 PT developing prods. for studying calcium channels, e.g. for
 PT obtaining agonists and antagonists
 PS Disclosure; Page 126-127; 285pp; English.
 CC The alpha 1D subunit cDNA has been isolated using fragments of the
 CC rabbit skeletal muscle calcium channel alpha 1 subunit cDNA as a
 CC probe to screen a cDNA library of human neuroblastoma cell line
 CC IMR32, to obtain clone alpha1.36. This clone was used as a probe to
 CC screen additional IMR32 cell cDNA libraries to obtain overlapping
 CC clones, which were then employed for screening until a sufficient
 CC series of clones to span the length of the nt sequence encoding the
 CC human alpha 1D subunit was obtd. Full-length clones were then
 CC constructed by ligating partial clones. Q84653 shows the nt sequence
 CC of the cDNA encoding the alpha 1D subunit. The Alpha 1D protein has
 CC a calculated Mr of 245,163. It contains four putative transmembrane
 CC repeated sequence regions which represent 24 putative transmembrane
 CC segments. It mediates DHP-sensitive high-voltage, long-lasting
 CC calcium channel activity. Q84654 shows an alternative exon encoding
 CC the 156 transmembrane domain. The difference occurs in Aa 373-
 CC 406.
 S0 Sequence 2161 Aa;
 Query Match 39.7%; Score 60; DB 14; Length 2161;
 Best Local Similarity 42.1%; Pred. No. 1.53e+01;
 Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 Db 989 lrp1rainrakgkhvqc 1007
 |||:::| | | : |
 Oy 2 LRPFKALVREKGRPSHSC 20

RESULT 10
 ID R71001 standard; Protein; 2161 AA.
 AC R71001;
 DT 30-NOV-1995 (first entry).
 DE Human neuronal calcium channel subunit alpha 1D.
 KM Calcium channel subunit; antagonist; agonist; diagnosis;
 KM Lambert Eaton Syndrome.
 OS Homo sapiens.
 PN M09504822-A.
 PD 16-FEB-1995.
 PF 11-AUG-1994; U09230.
 PR 11-AUG-1993; US-105536.
 PR 05-NOV-1993; US-149097.
 PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
 PI Ellis SB, Gillespie A, Harpold WM, Mccue AF, Williams ME;
 DR N-PSDB; 084653.
 DR N-PSDB; 084653.
 PT DNA encoding human calcium channel sub-unit(s) - used for
 PT developing prods. for studying calcium channels, e.g. for

PT obtaining agonists and antagonists
 PS Disclosure; Page 116-126; 285pp; English.
 CC The alpha 1D subunit cDNA has been isolated using fragments of the
 CC rabbit skeletal muscle calcium channel alpha 1 subunit cDNA as a
 CC probe to screen a cDNA library of human neuroblastoma cell line
 CC IMR32, to obtain clone alpha1.36. This clone was used as a probe to
 CC screen additional IMR32 cell cDNA libraries to obtain overlapping
 CC clones, which were then employed for screening until a sufficient
 CC series of clones to span the length of the nt sequence encoding the
 CC human alpha 1D subunit was obtd. Full-length clones were then
 CC constructed by ligating partial clones. Q84653 shows the nt sequence
 CC of the cDNA encoding the alpha 1D subunit. The Alpha 1D protein has
 CC a calculated Mr of 245,163. It contains four putative internal
 CC repeated sequence regions which represent 24 putative transmembrane
 CC segments. It mediates DHP-sensitive high-voltage, long-lasting
 CC calcium channel activity.
 S0 Sequence 2161 Aa;
 Query Match 39.7%; Score 60; DB 6; Length 2161;
 Best Local Similarity 42.1%; Pred. No. 1.53e+01;
 Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 Db 989 lrp1rainrakgkhvqc 1007
 |||:::| | | : |
 Oy 2 LRPFKALVREKGRPSHSC 20

RESULT 11
 ID R33545 standard; Protein; 2161 AA.
 AC R33545;
 DT 30-JUN-1993 (first entry)
 DE Sequence of the alpha 1D human calcium channel subunit.
 KM Human calcium channel subunit; diagnosis; agonist; antagonist;
 KM Lambert Eaton syndrome.
 OS Homo sapiens.
 PN M09304083-A.
 PD 04-MAR-1993.
 PF 14-AUG-1992; U06903.
 PR 15-AUG-1991; US-745206.
 PR 10-APR-1992; US-868354.
 PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
 PI Brenner R, Ellis SB, Feldman DH, Harpold WM, Mccue AF,
 PI Williams ME;
 DR WPI; 93-093936/11.
 DR N-PSDB; Q37811.
 PT DNA encoding specific human calcium channel sub-units - used for
 PT identifying calcium channel agonists and antagonists and
 PT diagnosing Lambert Eaton syndrome
 PS Disclosure; Page 93-101; 150pp; English.
 CC The alpha 1D subunit cDNA was isolated using fragments of the
 CC rabbit skeletal muscle calcium channel alpha 1 subunit cDNA as a
 CC probe to screen a cDNA library of a human neuroblastoma cell line,
 CC IMR32, to obtain clone alpha 1.36. This clone was used as a probe to
 CC screen additional IMR32 cell cDNA libraries to obtain overlapping
 CC clones which were then employed for screening until a sufficient
 CC series of clones to span the length of the nucleotide sequence
 CC encoding the human alpha 1D subunit were obtained, see Q37811.
 CC Q37812 provides the sequence of an alternative exon encoding the 156
 CC transmembrane domain of the alpha 1D subunit. The alpha 1D protein
 CC has a calculated Mr of 245,163.
 S0 Sequence 2161 Aa;
 Query Match 39.7%; Score 60; DB 6; Length 2161;
 Best Local Similarity 42.1%; Pred. No. 1.53e+01;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 989 lrp1rainrakgkhvqc 1007

Qy 2 LRPFKALVREKGRHSNC 20

RESULT 12

ID R71003 standard; Protein; 2163 AA.

AC R71003;

DT 30-NOV-1995 (first entry)

DE Human neuronal calcium channel subunit alpha 1c-1.

KM Calcium channel subunit; antagonist; agonist; diagnosis;

KM Lambert Eaton Syndrome.

OS Homo sapiens.

PN M09504822-A.

PD 16-FEB-1995.

PF 11-AUG-1994; U09230.

PR 11-AUG-1993; US-105536.

PR 05-NOV-1993; US-149097.

PA (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.

PI Ellis SB, Gillespie A, Harpold MM, McCue AF, Williams ME;

DR WPI; 95-090900/12.

DR N-PSDB; Q84655.

PT DNA encoding human calcium channel sub-unit(s) - used for

PT developing prods. for studying calcium channels, e.g. for

PT obtaining agonists and antagonists

PS Disclosure; Page 127-137; 285pp; English.

CC Numerous alpha 1c-specific cDNA clones were isolated in order to

CC characterise the alpha 1c coding sequence, the initiation of

CC translation and an alternatively spliced region. Q84655 sets

CC forth one alpha 1c coding sequence (alpha 1c-1) and R71003 sets

CC out its deduced AA sequence. Q87834 and R72607 set out another

CC splice variant, designated alpha 1c-2. Q84656 encodes an

CC alternative exon for the IV S1 transmembrane domain. Other

CC alpha 1c variants can be constructed by selecting alternative

CC amino terminal ends in place of the ends in Q84655 and Q87834

CC and/or inserting the alternative exon in the appropriate

CC location (see Q84655 FT). In addition, a nt. sequence (see Q84655

CC FT) can be deleted or inserted to produce an alternative alpha 1c

CC splice variant.

CC Sequence 2163 AA;

Query Match

Best Local Similarity 42.1%; Pred. No. 1.53e+01;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 1028 lrp1rainrakgkhvqc 1046

Qy 2 LRPFKALVREKGRHSNC 20

RESULT 13

ID R20784 standard; Protein; 349 AA.

AC R20784;

DT 12-MAY-1992 (first entry)

DE Tulip allylacyl amidase.

KM Chemical selection; herbicide resistance.

OS Tulippa.

PN J03290190-A.

PD 19-DEC-1991.

PF 06-APR-1990; 091565.

PR 06-APR-1990; JP-091565.

PA (NORQ) NORINSHO.

DR WPI; 92-045622/06.

DR N-PSDB; Q20497.

PT DNA encoding tulip allylacyl amidase - useful in chemical selection and

PT prodn. of herbicide resistant plants.

PS Disclosure; Fig 3; 11pp; Japanese.

CC Crude allylacyl amidase was prepd. from tulip bulbs and purified

CC by dialysis, DEAE cellulose chromatography, gel filtration and

CC hydroxyapatite chromatography. The purified protein was partially

CC digested by lysyl endopeptidase and the protein fragment sept. by

CC SDS-PAGE. The amino acid sequence was determined and the crude

CC protein used to prepare rabbit anti-allylacyl amidase serum to

CC screen a cDNA library, constructed in lambda gt11 using tulip bulb

CC poly A + RNA. The base sequence and corresp. amino acid sequence

CC were then determined.

CC Sequence 349 AA;

Query Match

Best Local Similarity 40.0%; Pred. No. 3.21e+01;

Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 5 kdlvgtvpphac 19

Qy 6 KALVREKGRHSNC 20

RESULT 14

ID R13944 standard; Protein; 553 AA.

AC R13944;

DT 27-NOV-1991 (first entry)

DE Partial HVT ribonucleotide reductase large subunit R1.

KM Newcastle disease; recombinant poultry virus vaccine.

OS Turkey herpes virus strain FC 126.

PN EP-447303-A.

PD 18-SEP-1991.

PF 07-MAR-1991; 400634.

PR 12-MAR-1990; FR-003105.

PA (RHON-) RHONE MERIEUX SA.

PI Rey-SeneLange A, Kohen G;

DR WPI; 91-275886/38.

DR N-PSDB; Q13430.

PT New herpes recombinants - useful as vaccines against human and

PT animal viral conditions, e.g. fowl pest, coccidiosis and

PT pasteurellosis

PS Disclosure; Page 13-14; 22pp; French.

CC This sequence is the C-terminal region of HVT R1. It was deduced

CC from a genomic DNA clone which was isolated because it contained

CC the entire R1 gene. Knowledge of the R1 small sub-unit sequence is

CC exploited in the construction of recombinant vaccines. A heterologous

CC sequence coding for an appropriate immunogen to protect against e.g.

CC Newcastle disease, avian anaemia, colibacillosis, avian infectious

CC bronchitis, etc. is inserted into the R1 gene. The HVT is suitable

CC for use in live vaccines as it is apathogenic and non-oncogenic.

CC See also Q13668.

CC Sequence 553 AA;

Query Match

Best Local Similarity 50.0%; Pred. No. 3.21e+01;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 415 lrpysqlrd 424

Qy 2 LRPFKALVRE 11

RESULT 15

ID R94346 standard; Protein; 3164 AA.

AC R94346;

DT 02-JUL-1996 (first entry)

DE Hepatitis GB virus (HGBV) clone GB contig A protein prod.

KM Hepatitis GB virus (HGBV); diagnosis; treatment; vaccine;

KM reagents; non-A; non-B; non-C; non-D; non-E; clone; GB contig A;

KM tamarin; infected plasma; lambda phage; cDNA library.

OS Hepatitis GB virus.

FH Key Location/Qualifiers

FT Misc difference 1..3164

FT /note="Others correspond to degenerate or STOP

FT codons in T00129"

PN M09521922-A2.

PD 17-AUG-1995.

PF 14-FEB-1995; U02118.

PR 14-FEB-1994; US-196030.

PR 13-MAY-1994; US-242654.

PR 29-JUL-1994; US-283314.

PR 23-NOV-1994; US-344190.

PR 23-NOV-1994; US-344185.

PR 27-JAN-1995; US-344557.

PA (AB80) ABBOTT LAB.

PI Bulik SL, Dawson GJ, Desai SM, Etker JC, Leary TP;

PI Muerthoff AS, Mushahwar IK, Pilot-Matias TJ, Schlauder GJ;

PI Simons JN;

DR WPI; 95-293123/38.

DR N-PSDB; T00129.

PT Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful

PT for diagnosis and therapy of hepatitis GB virus

PS Example 9; Pages 343-357; 661pp; English.

CC Double stranded hepatitis GB virus (HGBV) DNA obtd. from HGBV

CC infected tamarin plasma, using standard procedures, was used to

CC prepare a lambda phage HGBV cDNA library. Clones were rescued

CC from the lambda phage, searched against a sequence database and

CC found to be unique HGBV sequences. The clones were then used to

CC assemble the sequences T00129/30 (GB contig A and B) which encode

CC the proteins R94345-47 (the 3 possible coding strand reading

CC frames) and R82072, respectively. Reagents which comprise the HGBV

CC DNA, or its protein prode. can be used for the diagnosis, therapy

CC or in a vaccine to prevent HGBV infection.

SQ Sequence 3164 AA;

Query Match 37.7%; Score 57; DB 16; Length 3164;

Best Local Similarity 54.5%; Pred.No. 3.21e+01;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 1980 rrrgyrprhac 1990

Qy 10 REXGHRSHSC 20

Release 2.1D John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

(TM)

Merch_gp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Mar 24 11:18:31 1997; Maspar time 1.95 Seconds
Tabular output not generated.

Title: >US-08-644-289-1
Description: (1-20) from US08644289.pep
Perfect Score: 151
Sequence: 1 SLRPFKALVREKGRHSNC 20

Scoring table:
PAM 150
Gap 15

Searched: 52205 seqs, 18531385 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot33
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10

Statistics: Mean 28.995; Variance 36.679; scale 0.791

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Match	Length	DB	ID	Description	Pred. No.
No.	Score						
1	64	42.4	1166	1	ADDB_BACSU	ATP-DEPENDENT NUCLEAS	2.47e-01
2	62	41.1	164	9	X16_HUMAN	PRE-MRNA SPLICING FAC	6.13e-01
3	61	40.4	231	8	RM05_PROMI	MITOCHONDRIAL 60S RIB	9.57e-01
4	60	39.7	247	4	GRAB_MOUSE	GRANZYME B(G,H) PRECU	1.49e+00
5	60	39.7	256	10	YRDA_ECOLI	HYPOTHETICAL 28.4 KD	1.49e+00
6	60	39.7	466	4	HEMZ_ARATH	FERROCHELATASE, CHLO	1.49e+00
7	60	39.7	1852	2	C1C1_CYPCA	DIHYDROPYRIDINE-SENS	1.49e+00
8	60	39.7	1873	2	C1C1_RABIT	DIHYDROPYRIDINE-SENS	1.49e+00
9	60	39.7	2171	2	C1C1_RABIT	DIHYDROPYRIDINE-SENS	1.49e+00
10	59	39.1	299	10	YBBE_YEAST	POTATIVE MITOCHONDRIA	2.30e+00
11	59	39.1	392	6	NOIC_RHIFR	NOIC PROTEIN.	2.30e+00
12	58	38.4	531	3	G1IA_ORYSA	PROTEIN KINASE G1IA (3.53e+00
13	58	38.4	578	5	KPK7_ARATH	POTATIVE SERINE/THREO	3.53e+00

14	58	38.4	609	5	KXK1_PHAUV	PROTEIN KINASE PVK-1	3.53e+00
15	58	38.4	1157	7	POL_SFV3L	POL POLYPROTEIN (CONT	3.53e+00
16	57	37.7	193	9	VIF_SIVCZ	VIRION INFECTIVITY FA	5.39e+00
17	57	37.7	336	7	PYRD_ECOLI	DIHYDROROATE DEHYDR	5.39e+00
18	57	37.7	502	9	MASP_HUMAN	WISKOTT-ALDRICH SYND	5.39e+00
19	57	37.7	610	5	LIN9_CAEEL	LIN-9 PROTEIN.	5.39e+00
20	57	37.7	922	1	B3AT_CHICK	BAND 3 ANION TRANSPOR	5.39e+00
21	56	37.1	451	7	PHOQ_SALTY	VIRULENCE SENSOR PROT	8.18e+00
22	56	37.1	1234	1	B3A2_RAT	ANION EXCHANGE PROT	8.18e+00
23	56	37.1	1237	1	B3A2_RABIT	ANION EXCHANGE PROT	8.18e+00
24	56	37.1	1237	1	B3A2_MOUSE	ANION EXCHANGE PROT	8.18e+00
25	55	36.4	134	4	H3_ENTHI	HISTONE H3.	1.23e+01
26	55	36.4	328	10	Y1AO_HAEIN	HYPOTHETICAL PROTEIN	1.23e+01
27	55	36.4	364	1	ALFB_CHICK	FRUCTOSE-BISPHOSPHATE	1.23e+01
28	55	36.4	386	8	SENR_RAT	PROBABLE G PROTEIN-CO	1.23e+01
29	55	36.4	386	4	GPBE_RAT	PROBABLE G PROTEIN-CO	1.23e+01
30	55	36.4	436	1	B1OA_WJCLE	ADENOSYLMETHIONINE-8-	1.23e+01
31	55	36.4	586	9	VOLD_BBP2	OVERCOMING LYSGENIZA	1.23e+01
32	55	36.4	828	9	V2A_TAV	2A PROTEIN (PROBABLE	1.23e+01
33	55	36.4	879	5	LDLR_RAT	LOW-DENSITY LIPOPROTE	1.23e+01
34	55	36.4	988	9	TRA_FSENE	TRANSPORASE FOR TRANS	1.23e+01
35	55	36.4	1035	6	MT10_YEAST	SULFITE REDUCTASE (NA	1.23e+01
36	55	36.4	1238	1	BVGS_BORBR	VIRULENCE SENSOR PROT	1.23e+01
37	55	36.4	1238	1	BVGS_BORBR	VIRULENCE SENSOR PROT	1.23e+01
38	54	35.8	202	8	R84_EPIVI	CHLOROPLAST 30S RIBOS	1.85e+01
39	54	35.8	239	10	YGGH_ECOLI	HYPOTHETICAL 27.3 KD	1.85e+01
40	54	35.8	452	4	IDHP_HUMAN	ISOCITRATE DEHYDROGEN	1.85e+01
41	54	35.8	605	2	CACR_HUMAN	CARNITINE O-ACETYLTRA	1.85e+01
42	54	35.8	811	3	GAG_FOAMV	GAG POLYPROTEIN (CORE	1.85e+01
43	54	35.8	885	9	VGLB_HSV2S	GLYCOPROTEIN B PRECUR	1.85e+01
44	54	35.8	2131	2	C1NA_THROM	SODIUM CHANNEL PROTEI	1.85e+01
45	54	35.8	2813	9	VWF_HUMAN	VON WILLEBRAND FACTOR	1.85e+01

ALIGNMENTS

RESULT	1	STANDARD;	PRT;	1166 AA.
ID	ADDB_BACSU			
AC	P23477;			
DT	01-NOV-1991 (REL. 20, CREATED)			
DT	01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
DE	ATP-DEPENDENT NUCLEASE SUBUNIT B.			
GN	ADDB.			
OS	BACILLUS SUBTILIS.			
OC	PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-OG1;			
RX	MEDLINE; 91267926.			
RA	KOOLISTRA J., VENEMA G.,			
RL	J. BACTERIOL. 173:3644-3655(1991).			
CC	-1- FUNCTION: THE ENZYME COMPLEX MAY HAVE A WIDE VARIETY OF CATALYTIC			
CC	ACTIVITIES INCLUDING ATP-DEPENDENT EXONUCLEASE, ATP-STIMULATED			
CC	ENDONUCLEASE, ATP-DEPENDENT HELICASE AND DNA-DEPENDENT ATPASE			
CC	ACTIVITIES.			
CC	-1- SUBUNIT: THE B.SUBTILIS ATP-DEPENDENT NUCLEASE COMPLEX IS FORMED			
CC	BY ONE B SUBUNIT IN CONJUNCTION WITH ONE A SUBUNIT.			
DR	EMBL; M63489; G142439; -.			
DR	PIR; A39432; A39432.			
DR	SUBTILIST; BG10465; ADDB.			
KM	HYDROLASE; NUCLEASE; EXONUCLEASE; HELICASE;			
KM	ATP-BINDING; DNA REPAIR.			
FT	NP_BIND	1	22	ATP (POTENTIAL).

SQ SEQUENCE 1166 AA; 134631 MM; 91EEB84F CRC32;

Query Match 42.4%; Score 64; DB 1; Length 1166;
 Best Local Similarity 69.2%; Pred. No. 2.47e-01;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 329 reihsalvrekgyr 341

Qy 3 REKALVREKGRH 15

RESULT 2
 ID X16 HUMAN STANDARD; PRT; 164 AA.
 AC P23152;
 DT 01-NOV-1991 (REL. 20, CREATED)
 DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE PRE-MRNA SPLICING FACTOR SRP20 (X16 PROTEIN).
 GN X16.
 OS HOMO SAPIENS (HUMAN), AND MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUETHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=HUMAN;
 RX MEDLINE; 92249775.
 RA ZAHLER A.M., LANE W.S., STOLK J.A., ROTH M.B.;
 RL GENES DEV. 6:837-847(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=MOUSE;
 RX MEDLINE; 91232908.
 RA AYANE M., PREUSS U., KOEHLER G., NIELSEN P.J.;
 RL NUCLEIC ACIDS RES. 19:1273-1278(1991)
 CC -1- FUNCTION: MAY BE INVOLVED IN RNA PROCESSING IN RELATION WITH
 CC CELLULAR PROLIFERATION AND/OR MATURATION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN THYMUS AND PRE-B CELL
 CC LINES; HIGH IN TESTIS, BRAIN AND SPLEEN; VERY LOW IN HEART AND
 CC NOT DETECTABLE IN LIVER AND KIDNEY.
 CC -1- INDUCTION: BY SERUM, IN A TISSUE CULTURE.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNP).
 DR EMBL; L10838; G338484; -.
 DR EMBL; X53824; G55440; -.
 DR PIR; S14016; S14016.
 DR PROSITE; PS00030; RNP 1.
 KM NUCLEAR PROTEIN; RNA-BINDING; mRNA SPLICING; REPEAT.
 FT DOMAIN 46 53 RNA-BINDING (RNP1) (BY SIMILARITY).
 FT DOMAIN 105 164 ARG/LYS-RICH (HIGHLY BASIC; MAY BE
 FT REPEAT 119 133 INVOLVED IN NUCLEIC ACIDS BINDING).
 FT REPEAT 149 164 2 X APPROXIMATE REPEATS, BASIC.
 FT SIMILAR 105 164 B-1.
 FT SIMILAR 164 AA; 19329 MM; 531E77C2 CRC32; B-2.
 SQ SEQUENCE 164 AA; 19329 MM; 531E77C2 CRC32; TO PROTEIN.

Query Match 41.1%; Score 62; DB 9; Length 164;
 Best Local Similarity 50.0%; Pred. No. 6.13e-01;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 137 rslsternhkpars 150
 Qy 6 KALVREKGRHS 19

RESULT 3
ID RM05 PROM1 STANDARD; PRT; 231 AA.

AC P46749;

DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE MITOCHONDRIAL 60S RIBOSOMAL PROTEIN L5.

GN RPL5.

OS PROTOCECA WICKERHAMII.

OC MITOCHONDRION.

OC EUKARYOTA; PLANTA; PHYCOPHYTA; CHLOROPHYTA (GREEN ALGAE);

OC CHLOROPHYCEAE; CHLOROCOCCALES; OOCYSTACEAE.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=263-11;

RX MEDLINE; 94180393.

RA WOLFE G., PLANTE T., LANG B.F., KUECK U., BORGER G.;

RL J. MOL. BIOL. 237:75-86(1994).

CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.

CC -1- SIMILARITY: BELONGS TO THE L5P FAMILY OF RIBOSOMAL PROTEINS.

DR EMBL; U02970; G467861; -.

DR PROSITE; PS00358; RIBOSOMAL L5.

KM RIBOSOMAL PROTEIN; MITOCHONDRION.

SQ SEQUENCE 231 AA; 26584 MM; E827E483 CRC32;

Query Match 40.4%; Score 61; DB 8; Length 231;

Best Local Similarity 36.8%; Pred. No. 9.57e-01;

Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Db 88 lpfatlamtdgkpytc 106

Qy 2 LRFKALVREKGRHS 20

RESULT 4
 ID GRAB MOUSE STANDARD; PRT; 247 AA.
 AC P04187;
 DT 20-MAR-1987 (REL. 04, CREATED)
 DT 20-MAR-1987 (REL. 04, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE GRABZYME B(G,H) PRECURSOR (EC 3.4.21.79) (CYTOTOXIC CELL PROTEASE 1)
 DE (CCP1) (CTLA-1) (FRAGMENTIN 2).
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUETHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86208120.
 RA LOBE C.G., FINLAY B.B., PARANCHYCH W., PAETZAU V.H., BLECKLEY R.C.;
 RL SCIENCE 232:858-861(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89062424.
 RA LOBE C.G., UPTON C., DUGGAN B., EHRMAN N., LETELLIER M., BELL J.,
 RA MCFADDEN G., BLECKLEY R.C.;
 RL BIOCHEMISTRY 27:6941-6946(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86284960.
 RA BRUNET J.F., DOSSETO M., DENIZOT F., MATTEI M.G., CLARK W.R.,
 RA HAQOI T.M., FERRIER P., NABHOLZ M., SCHMITT-VERHULST A.M.,
 RA LUCIANI M.F., GOLSTEIN P.;
 RL NATURE 322:268-271(1986).

RN (4)
 RP SEQUENCE OF 21-40.
 RX MEDLINE: 87215932.
 RA MASSON D., TSCHOPP J.;
 RL CELL 49:679-685 (1987).
 RN [5]
 RP 3D-STRUCTURE MODELLING.
 RX MEDLINE: 89184501.
 RA MURPHY M.E.P., MOULT J., BLACKLEY R.C., GERSHENFELD H.,
 RL WEISSMAN I.L., JAMES M.N.G.;
 CC PROTEINS 4:190-204 (1988).
 CC -1- FUNCTION: THIS ENZYME IS NECESSARY FOR TARGET CELL LYSIS IN CELL-
 CC MEDIATED IMMUNE RESPONSES. IT CLEAVES AFTER ASP. MAY BE INVOLVED
 CC IN APOPTOSIS.
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ASP-1-XAA >> ASN-1-XAA
 CC > MET-1-XAA, SER-1-XAA.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES OF CYTOLYTIC
 CC T-LYMPHOCYTES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY. STRONGEST TO OTHER GRANZYMES AND TO MAST CELL
 CC PROTEASES.
 DR EMBL; X04072; G50587; -.
 DR EMBL; M12302; G309154; -.
 DR EMBL; M22526; G201027; -.
 DR PIR; A00956; PRMSCL.
 DR PIR; A28952; A28952.
 DR PIR; B26944; B26944.
 DR PDB; 2CP1; 15-OCT-94.
 DR PROSITE; PS00134; TRYPSIN HIS.
 DR PROSITE; PS00135; TRYPSIN SER.
 KM HYDROLASE; SERINE PROTEASE; ZIMOGEN; SIGNAL; T-CELL; CYTOLYSIS;
 KM APOPTOSIS; 3D-STRUCTURE.
 FT SIGNAL 1 18
 FT PROPER 19 20 ACTIVATION PEPTIDE.
 FT CHAIN 21 247 GRANZYME B.
 FT ACT SITE 64 64 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 108 108 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 203 203 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 49 65 BY SIMILARITY.
 FT DISULFID 142 209 BY SIMILARITY.
 FT DISULFID 173 188 BY SIMILARITY.
 FT CARBOHYD 71 71 POTENTIAL.
 FT CARBOHYD 182 182 POTENTIAL.
 SQ SEQUENCE 247 AA; 27470 MW; E05916CB CRC32;
 Query Match 39.7%; Score 60; DB 4; Length 247;
 Best Local Similarity 38.9%; Pred. No. 1.49e+00;
 Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
 Db 32 rymalleikdgqpaic 49
 Qy 3 RPKALVREKGRHSHSC 20
 RESULT 5
 ID YRDA ECOLI STANDARD; PRT; 256 AA.
 AC P45770;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 28.4 KD PROTEIN IN RND-AOE INTERGENIC REGION (0256).
 GN YRDA.
 OS ESCHERICHIA COLI.
 OC PROKARYOTA; FRACILLICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;

OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RA PLUNKETT G. III;
 RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- SIMILARITY: TO P.AERUGINOSA FERRIPYOCHELIN BINDING PROTEIN AND
 CC E. COLI PROTEIN CALE.
 DR EMBL; U18997; G606213; -.
 DR ECOGENE; EGI2838; YRDA.
 KM HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
 FT TRANSMEM 13 33 POTENTIAL.
 SQ SEQUENCE 256 AA; 28375 MW; 725231B1 CRC32;
 Query Match 39.7%; Score 60; DB 10; Length 256;
 Best Local Similarity 42.9%; Pred. No. 1.49e+00;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Db 77 lrpvrdlfpqgr 90
 Qy 2 LRPRALVREKGRH 15
 RESULT 6
 ID HEMZ ARATH STANDARD; PRT; 466 AA.
 AC P42043;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE FERROCHELATASE, CHLOROPLAST PRECURSOR (EC 4.99.1.1) (PROTOHEME
 DE FERRO-LYASE) (HEME SYNTHETASE).
 GN HEM15.
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 OC CAPPARALE; CRUCIFERAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. LANDSBERG ERECTA; TISSUE-WHOLE SEEDLINGS;
 RX MEDLINE: 94230447.
 RA SMITH A.G., SANTANA M.A., WALLACE-COOK A.D.M., ROBER J.M.,
 RA LABBE-BOIS R.;
 RL J. BIOL. CHEM. 269:13405-13413 (1994).
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRIN + FE(2+) = PROTOHEME + 2 H(+).
 CC -1- PATHWAY: LAST STEP IN PROTOHEME BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- TISSUE SPECIFICITY: PRESENT IN BOTH LEAVES AND ROOTS.
 CC -1- SIMILARITY: TO OTHER SPECIES FERROCHELATASE.
 DR EMBL; X73417; G511081; -.
 DR PROSITE; PS00534; FERROCHELATASE.
 KM PORPHYRIN BIOSYNTHESIS; HEME BIOSYNTHESIS; LYASE; IRON; CHLOROPLAST;
 KM TRANSIT PEPTIDE.
 FT TRANSIT 1 2 CHLOROPLAST (POTENTIAL).
 FT CHAIN 1 466 FERROCHELATASE.
 SQ SEQUENCE 466 AA; 52032 MW; 2D3756C1 CRC32;
 Query Match 39.7%; Score 60; DB 4; Length 466;
 Best Local Similarity 43.8%; Pred. No. 1.49e+00;
 Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 Db 10 fplptktdhfrpsc 25
 Qy 5 FRALVREKGRHSHSC 20

Mar 24 11:16

US-08-644-289-1.rsp

7

RESULT 7
ID C1C1 CYPEA STANDARD; PRT; 1852 AA.
AC P22316;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE DIHYDROPIRIDINE-SENSITIVE L-TYPE, SKELETAL MUSCLE CALCIUM CHANNEL
DE ALPHA-1 SUBUNIT.
OS CYPRINUS CARPIO (COMMON CARP).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNAETHOSTOMATA;
OC OSTEICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE;
RX MEDLINE; 91126068.
RA GRABNER M., FRIEDRICH K., KNAUS H.-G., STRIESSNIG J., SCHEFFAUER F.,
RA STRANDINGER R., KOCH W.J., SCHWARTZ A., GLOSSMANN H.;
RL PROC. NATL. ACAD. SCI. U.S.A. 88:727-731(1991).
CC -1- FUNCTION: CALCIUM CHANNEL PROTEIN WHICH PLAYS AN IMPORTANT ROLE IN
EXCITATION-CONTRACTION COUPLING. THE ALPHA-1 SUBUNIT BINDS
DIHYDROPIRIDINE (DHP), A CALCIUM CHANNEL BLOCKER.
CC -1- SUBUNIT: THIS L-TYPE CALCIUM CHANNEL IS COMPOSED OF TWO SUBUNITS:
ALPHA-1, AND ALPHA-2.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- PTM: MAY NOT BE PHOSPHORYLATED.
CC -1- TISSUE SPECIFICITY: SKELETAL MUSCLE.
CC -1- SIMILARITY: TO OTHER CALCIUM CHANNEL ALPHA SUBUNITS.
DR EMBL; M62554; G213050; -.
DR PIR; A31860; A37860.
KM IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL;
KM CALCIUM CHANNEL; GLYCOPROTEIN; DUPLICATION; PHOSPHORYLATION.
FT REPEAT 57 354 I.
FT REPEAT 433 679 II.
FT REPEAT 845 1084 III.
FT REPEAT 1121 1405 IV.
FT TRANSMEM 71 86 S1 OF REPEAT I.
FT TRANSMEM 108 127 S2 OF REPEAT I.
FT TRANSMEM 140 155 S3 OF REPEAT I.
FT TRANSMEM 177 195 S4 OF REPEAT I.
FT TRANSMEM 215 234 S5 OF REPEAT I.
FT TRANSMEM 327 351 S6 OF REPEAT I.
FT TRANSMEM 448 466 S1 OF REPEAT II.
FT TRANSMEM 482 501 S2 OF REPEAT II.
FT TRANSMEM 510 528 S3 OF REPEAT II.
FT TRANSMEM 539 557 S4 OF REPEAT II.
FT TRANSMEM 577 596 S5 OF REPEAT II.
FT TRANSMEM 652 675 S6 OF REPEAT II.
FT TRANSMEM 816 834 S1 OF REPEAT III.
FT TRANSMEM 851 870 S2 OF REPEAT III.
FT TRANSMEM 883 901 S3 OF REPEAT III.
FT TRANSMEM 909 927 S4 OF REPEAT III.
FT TRANSMEM 947 966 S5 OF REPEAT III.
FT TRANSMEM 1057 1081 S6 OF REPEAT III.
FT TRANSMEM 1135 1153 S1 OF REPEAT IV.
FT TRANSMEM 1169 1188 S2 OF REPEAT IV.
FT TRANSMEM 1197 1215 S3 OF REPEAT IV.
FT TRANSMEM 1253 1271 S4 OF REPEAT IV.
FT TRANSMEM 1291 1310 S5 OF REPEAT IV.
FT TRANSMEM 1378 1402 S6 OF REPEAT IV.
FT CARBOHYD 99 99 POTENTIAL.
FT CARBOHYD 102 102 POTENTIAL.
FT CARBOHYD 274 274 POTENTIAL.
FT CARBOHYD 470 470 POTENTIAL.
FT CARBOHYD 1157 1157 POTENTIAL.

Mar 24 11:16

US-08-644-289-1.rsp

8

FT MOD_RES 407 407 PHOSPHORYLATION (BY CARP)
FT (POTENTIAL).
FT MOD_RES 1471 1471 PHOSPHORYLATION (BY CARP)
FT (POTENTIAL).
FT MOD_RES 1523 1523 PHOSPHORYLATION (BY CARP)
FT (POTENTIAL).
FT MOD_RES 1738 1738 PHOSPHORYLATION (BY CARP)
FT (POTENTIAL).
SQ SEQUENCE 1852 AA; 210096 MW; 46513952 CRC32;
Query Match 39.7%; Score 60; DB 2; Length 1852;
Best Local Similarity 42.1%; Pred. No. 1.49e+00;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
Db 918 lrp1rainrakg1khvqc 936
|||:::| | | | |
Qy 2 lrp1rainrakg1khvqc 20
RESULT 8
ID C1C1 RABIT STANDARD; PRT; 1873 AA.
AC P07293;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE DIHYDROPIRIDINE-SENSITIVE L-TYPE, SKELETAL MUSCLE CALCIUM CHANNEL
DE ALPHA-1 SUBUNIT.
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; LAGOMORPHA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE;
RX MEDLINE; 87258269.
RA TANABE T., TAKEISHIMA H., MIKAMI A., FLOCKERZI V., TAKAHASHI H.,
RA KANGAMA K., KOJIMA M., MATSUO H., HIROSE T., NIMA S.;
RL NATURE 328:313-318(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE;
RX MEDLINE; 88336904.
RA ELLIS S.B., WILLIAMS M.E., WAYS N.R., BRENNER R., SHARP A.H.,
RA LEUNG A.T., CAMPBELL K.P., MCKENNA E., KOCH W.J., HUI A., SCHWARTZ A.,
RA HARPOLD M.M.;
RL SCIENCE 241:1661-1664(1988).
CC -1- FUNCTION: CALCIUM CHANNEL PROTEIN WHICH PLAYS AN IMPORTANT ROLE IN
EXCITATION-CONTRACTION COUPLING. THE ALPHA-1 SUBUNIT BINDS
DIHYDROPIRIDINE (DHP), A CALCIUM CHANNEL BLOCKER.
CC -1- SUBUNIT: THE L-TYPE CALCIUM CHANNEL IS COMPOSED OF FOUR SUBUNITS:
ALPHA-1, ALPHA-2, BETA AND GAMMA.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: SKELETAL MUSCLE.
CC -1- SIMILARITY: TO OTHER CALCIUM CHANNEL ALPHA SUBUNITS.
DR EMBL; X05921; G1548; -.
DR EMBL; M23919; G164761; -.
DR PIR; A30063; A30063.
KM IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL;
KM CALCIUM CHANNEL; GLYCOPROTEIN; DUPLICATION; PHOSPHORYLATION.
FT REPEAT 38 337 I.
FT REPEAT 418 664 II.
FT REPEAT 786 1068 III.
FT REPEAT 1105 1384 IV.
FT TRANSMEM 52 70 S1 OF REPEAT I.
FT TRANSMEM 89 108 S2 OF REPEAT I.

Mar 24 11:16

US-08-644-289-1.rsp

9

FT TRANSMEM 121 139 S3 OF REPEAT I.
FT TRANSMEM 161 179 S4 OF REPEAT I.
FT TRANSMEM 199 218 S5 OF REPEAT I.
FT TRANSMEM 310 334 S6 OF REPEAT I.
FT TRANSMEM 433 451 S1 OF REPEAT II.
FT TRANSMEM 467 486 S2 OF REPEAT II.
FT TRANSMEM 495 513 S3 OF REPEAT II.
FT TRANSMEM 524 542 S4 OF REPEAT II.
FT TRANSMEM 562 581 S5 OF REPEAT II.
FT TRANSMEM 637 661 S6 OF REPEAT II.
FT TRANSMEM 800 818 S1 OF REPEAT III.
FT TRANSMEM 835 854 S2 OF REPEAT III.
FT TRANSMEM 867 885 S3 OF REPEAT III.
FT TRANSMEM 893 911 S4 OF REPEAT III.
FT TRANSMEM 931 950 S5 OF REPEAT III.
FT TRANSMEM 1041 1065 S6 OF REPEAT III.
FT TRANSMEM 1119 1137 S1 OF REPEAT IV.
FT TRANSMEM 1153 1172 S2 OF REPEAT IV.
FT TRANSMEM 1181 1199 S3 OF REPEAT IV.
FT TRANSMEM 1232 1250 S4 OF REPEAT IV.
FT TRANSMEM 1270 1289 S5 OF REPEAT IV.
FT TRANSMEM 1357 1381 S6 OF REPEAT IV.
FT CARBOHYD 79 79 POTENTIAL.
FT CARBOHYD 257 257 POTENTIAL.
FT MOD RES 687 687 PHOSPHORYLATION (BY CAPK).
FT CONFLICT 694 694 T -> R (IN REF. 2).
FT CONFLICT 1808 1808 T -> M (IN REF. 2).
FT CONFLICT 1815 1815 A -> V (IN REF. 2).
FT CONFLICT 1835 1835 A -> E (IN REF. 2).
SQ SEQUENCE 1873 AA; 212028 MW; BA27142C CRC32;

Query Match 39.7%; Score 60; DB 2; Length 1873;
Best Local Similarity 42.1%; Pred. No. 1.49e+00;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 902 Lrp1rainrakjkhvqc 920
||||:| | | : |
Qy 2 LRPFAIVREKGRPSHC 20

RESULT 9
ID CICC RABIT STANDARD; PRF; 2171 AA.
AC P153R1;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE DIHYDROPIRIDINE-SENSITIVE L-TYPE, CARDIAC MUSCLE CALCIUM CHANNEL
DE ALPHA-1 SUBUNIT.
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; LAGOMORPHA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RX MEDLINE; 89330539.
RA MIKAMI A., IMOTO K., TANABE T., NI IDOME T., MORI Y., TAKESHIMA H.,
RA NAKAMURA S., NUKA S.;
RL NATURE 340:230-233(1989).
CC -1- FUNCTION: CALCIUM CHANNEL PROTEIN WHICH PLAYS AN IMPORTANT ROLE IN
CC EXCITATION-CONTRACTION COUPLING. THE ALPHA-1 SUBUNIT BINDS
CC DIHYDROPIRIDINE (DHP), A CALCIUM CHANNEL BLOCKER.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: CARDIAC MUSCLE.
CC -1- SIMILARITY: TO OTHER CALCIUM CHANNEL ALPHA SUBUNITS.

Mar 24 11:16

US-08-644-289-1.rsp

10

DR EMBL; X15539; G1510; -.
DR PIR; S05054; S05054.
KM IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL;
KM CALCIUM CHANNEL; GLYCOPROTEIN; DUPLICATION; PHOSPHORYLATION.

FT REPEAT 141 438 I.
FT REPEAT 540 786 II.
FT REPEAT 917 1199 III.
FT REPEAT 1236 1509 IV.
FT TRANSMEM 155 173 S1 OF REPEAT I.
FT TRANSMEM 191 211 S2 OF REPEAT I.
FT TRANSMEM 224 242 S3 OF REPEAT I.
FT TRANSMEM 263 281 S4 OF REPEAT I.
FT TRANSMEM 301 320 S5 OF REPEAT I.
FT TRANSMEM 411 435 S6 OF REPEAT I.
FT TRANSMEM 555 573 S1 OF REPEAT II.
FT TRANSMEM 589 608 S2 OF REPEAT II.
FT TRANSMEM 617 635 S3 OF REPEAT II.
FT TRANSMEM 646 664 S4 OF REPEAT II.
FT TRANSMEM 684 703 S5 OF REPEAT II.
FT TRANSMEM 759 783 S6 OF REPEAT II.
FT TRANSMEM 931 949 S1 OF REPEAT III.
FT TRANSMEM 966 985 S2 OF REPEAT III.
FT TRANSMEM 998 1016 S3 OF REPEAT III.
FT TRANSMEM 1024 1042 S4 OF REPEAT III.
FT TRANSMEM 1062 1081 S5 OF REPEAT III.
FT TRANSMEM 1172 1196 S6 OF REPEAT III.
FT TRANSMEM 1250 1268 S1 OF REPEAT IV.
FT TRANSMEM 1284 1303 S2 OF REPEAT IV.
FT TRANSMEM 1312 1330 S3 OF REPEAT IV.
FT TRANSMEM 1355 1373 S4 OF REPEAT IV.
FT TRANSMEM 1393 1412 S5 OF REPEAT IV.
FT TRANSMEM 1482 1506 S6 OF REPEAT IV.
FT CARBOHYD 100 100 POTENTIAL.
FT CARBOHYD 183 183 POTENTIAL.
FT CARBOHYD 358 358 POTENTIAL.
FT CARBOHYD 493 493 POTENTIAL.
FT CARBOHYD 859 859 POTENTIAL.
FT CARBOHYD 928 928 POTENTIAL.
FT CARBOHYD 1247 1247 POTENTIAL.
FT CARBOHYD 1418 1418 POTENTIAL.
FT CARBOHYD 1469 1469 POTENTIAL.
SQ SEQUENCE 2171 AA; 242783 MW; 4FC5E64A CRC32;

Query Match 39.7%; Score 60; DB 2; Length 2171;
Best Local Similarity 42.1%; Pred. No. 1.49e+00;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 1033 Lrp1rainrakjkhvqc 1051
||||:| | | : |
Qy 2 LRPFAIVREKGRPSHC 20

RESULT 10
ID Y88E YEAST STANDARD; PRF; 299 AA.
AC P38152;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE PUTATIVE MITOCHONDRIAL CARRIER YBR291C.
GN YBR291C OR YBR2039.
OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCETES.
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=5288C;
RX MEDLINE; 94378722.
RA HOLMSTROM K., BRANDT T., KALLESOE T.;
RL YEAST 10:547-562(1994).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (POTENTIAL).
CC -|- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
DR EMBL; X76053; G429131; -.
DR EMBL; Z36160; G536746; -.
DR PIR; S44554; S44554.
DR PIR; S39148; S39148.
DR PROSITE; PS00215; MITOCH. CARRIER.
KM HYPOTHETICAL PROTEIN; MITOCHONDRION; INNER MEMBRANE; REPEAT;
KM TRANSMEMBRANE; TRANSPORT.
FT TRANSMEM 16 36 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 113 133 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
SQ SEQUENCE 299 AA; 32173 MW; 01FE9E27 CRC32;

Query Match
Best Local Similarity 39.1%; Score 59; DB 10; Length 299;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 156 vnyesylvrdkg 167
: : : ||| : ||
Qy 2 LRPFKALNREKG 13

RESULT 11
ID NCLOC_RHIFR STANDARD; PRT; 392 AA.
AC P26508;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE NCLOC PROTEIN.
GN NCLOC.
OS RHIZOBIUM FREDDII.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC RHIZOBIACEAE.
CC RHIZOBIUM
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA257;
RX MEDLINE; 91260457.
RA KRISHNAN H.B., PUEPKKE S.G.;
RL MOL. MICROBIOL. 5:737-745(1991).
CC -|- SIMILARITY: TO PROKARYOTIC DNAJ, AND TO EUKARYOTIC DNAJ-LIKE
CC PROTEINS.
DR EMBL; L03521; G152354; -.
DR PIR; S15295; S15295.
DR PROSITE; PS00636; DNAJ_NTER.
KM CHAPERONE; MODULATION.
FT DOMAIN 2 71 DNAJ-LIKE.
FT DOMAIN 76 93 GLY-RICH.
FT DOMAIN 96 121 ARC-RICH.
SQ SEQUENCE 392 AA; 43565 MW; CC03D880 CRC32;

Query Match
Best Local Similarity 39.1%; Score 59; DB 6; Length 392;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 151 lrrhnglgreaqhp 165
: : : | | | | | | | |

```

	Qy	2	LRFKALVREKGRR	16
RESULT	12			
ID	GILIA ORYSA	STANDARD;	PRT;	531 AA.
AC	p47997;			
DT	01-FEB-1996 (REL. 33,	CREATED)		
DT	01-FEB-1996 (REL. 33,	LAST SEQUENCE UPDATE)		
DT	01-FEB-1996 (REL. 33,	LAST ANNOTATION UPDATE)		
DE	PROTEIN KINASE GIIA (EC 2.7.1.-) (FRAGMENT).			
OS	ORYZA SATIVA (RICE).			
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;			
CC	CAPPALES; GRAMINEAE.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. IR36; TISSUE=LEAF;			
RX	MEDLINE; 89240692.			
RA	LAWTON M.A., YAMAMOTO R.T., HANKS S.K., LAMB C.J.;			
RL	PROC. NATL. ACAD. SCI. U.S.A. 86:3140-3144(1989).			
CC	-!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF METABOLISM AND SIGNAL TRANSDUCTION PROCESSES.			
CC	-!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-PROTEIN KINASES. STRONG, TO P.VULGARIS PVK-1.			
DR	EMBL; J04556; g169788; -.			
KW	TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING.			
FT	NON_TER	1		
FT	DOMAIN	142	480	PROTEIN KINASE.
FT	NP_BIND	148	156	ATP (BY SIMILARITY).
FT	BINDING	171	171	ATP (BY SIMILARITY).
FT	ACT_SITE	267	267	BY SIMILARITY.
FT	DOMAIN	317	339	CYS-RICH INSERT.
FT	NON_TER	531	531	
SQ	SEQUENCE	531 AA;	58647 MW;	FAF6904A CRC32;
Query Match		38.4%;	Score 58;	DB 3;
Best Local Similarity		61.5%;	Pred. No. 3.53e+00;	Length 531;
Matches	8;	Conservative	2;	Mismatches 3;
				Indels 0;
				Gaps 0;
D6	268 lkenvlyvredgh	280		
Qy	2 LRFKALVREKGRR	14		
RESULT	13			
ID	KPKT ARATH	STANDARD;	PRT;	578 AA.
AC	005999;			
DT	01-FEB-1995 (REL. 31,	CREATED)		
DT	01-FEB-1995 (REL. 31,	LAST SEQUENCE UPDATE)		
DT	01-NOV-1995 (REL. 32,	LAST ANNOTATION UPDATE)		
DE	PUTATIVE SERINE/THREONINE-PROTEIN KINASE PK7 (EC 2.7.1.-).			
OS	ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).			
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;			
CC	CAPPALES; CRUCIFERAE.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE; 93077048.			
RA	HAYASHIDA N., MIZOGUCHI T., YAMAGUCHI-SHINOZAKI K., SHINOZAKI K.;			
RL	GENE 121:325-330(1992).			
CC	-!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN ROOT TISSUE			
CC	WITH LOWER LEVELS FOUND IN LEAF, STEM, SEED AND FLOWER.			
CC	-!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-PROTEIN KINASES.			
CC	EMBL; S50281; -; NOT_ANNOTATED_CDS.			

DR PIR; JC1385; JC1385.
 DR PROSITE; P500107; PROTEIN KINASE ATP.
 DR PROSITE; P500108; PROTEIN KINASE ST.
 DR PROSITE; P500111; PROTEIN_KINASE_DOM.
 KM TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING.
 FT DOMAIN 182 516 PROTEIN KINASE.
 FT NP BIND 188 196 ATP (BY SIMILARITY).
 FT BINDING 211 211 ATP (BY SIMILARITY).
 FT ACT SITE 307 307 BY SIMILARITY.
 SQ SEQUENCE 578 AA; 64286 MW; 9DB8E278 CRC32;
 Query Match 38.4%; Score 58; DB 5; Length 578;
 Best Local Similarity 61.5%; Pred. No. 3.53e+00;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 308 lkpennlvredqh 320
 l:l : llll l
 Qy 2 LRPFKALVREKGH 14

RESULT 14
 ID KPKL PHAVU STANDARD; PRT; 609 AA.
 AC P15792;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE PROTEIN KINASE PYR-1 (EC 2.7.1.-).
 OS PHASEOLUS VULGARIS (KIDNEY BEAN) (FRENCH BEAN).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
 NC FABACEAE.
 RC [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89240692.
 RA LAWTON M.A., YAMAMOTO R.T., HANKS S.K., LAMB C.J.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:3140-3144(1989).
 CC -1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
 CC PROTEIN KINASES. STRONG, TO RICE G11A.
 DR EMBL; J04555; G169361; -.
 DR PIR; A30311; A30311.
 DR HSP; P05132; ICTP.
 DR PROSITE; P500107; PROTEIN KINASE ATP.
 DR PROSITE; P500108; PROTEIN KINASE ST.
 DR PROSITE; P500111; PROTEIN_KINASE_DOM.
 KM TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING.
 FT DOMAIN 229 565 PROTEIN KINASE.
 FT NP BIND 235 243 ATP (BY SIMILARITY).
 FT BINDING 258 258 ATP (BY SIMILARITY).
 FT ACT SITE 354 354 BY SIMILARITY.
 FT DOMAIN 400 421 CYS-RICH INSERT.
 SQ SEQUENCE 609 AA; 68101 MW; A7E5A35F CRC32;

Query Match 38.4%; Score 58; DB 5; Length 609;
 Best Local Similarity 61.5%; Pred. No. 3.53e+00;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 355 lkpennlvredqh 367
 l:l : llll l
 Qy 2 LRPFKALVREKGH 14

RESULT 15
 ID POL_SFV3L STANDARD; PRT; 1157 AA.
 AC P27401;
 DT 01-AUG-1992 (REL. 23, CREATED)

DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
 DE POL POLYPROTEIN (CONTAINS: PROTEASE (EC 3.4.23.-); REVERSE
 DE TRANSCRIPTASE (EC 2.7.7.49); ENDONUCLEASE).
 CN POL.
 OS SIMIAN FOAMY VIRUS (TYPE 3 / STRAIN LK3) (SFV-3).
 OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
 OC SPUMAVIRINAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92124734.
 RA RENNE R., FRIEDL E., SCHWEIZER M., FLEPS U., TURK R.,
 RA NEUMANN-HAEFELIN D.;
 RL VIROLOGY 186:597-608(1992).
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- THIS POLYPROTEIN IS PROBABLY SYNTHESIZED AS A GAG-POL POLYPROTEIN
 CC BY A +1 FRAMESHIFT.
 DR EMBL; M74895; G334872; -.
 DR PIR; B40820; GNLIJK.
 KM HYDROLASE; ASPARTYL PROTEASE; RNA-DIRECTED DNA POLYMERASE;
 KM ENDONUCLEASE; POLYPROTEIN.
 FT ACT SITE 38 38 BY SIMILARITY.
 SQ SEQUENCE 1157 AA; 131225 MW; C1425458 CRC32;

Query Match 38.4%; Score 58; DB 7; Length 1157;
 Best Local Similarity 44.4%; Pred. No. 3.53e+00;
 Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Db 726 lkpdlilhekqnpas 743
 l:l : llll l
 Qy 2 LRPFKALVREKGRPSHS 19

Search completed: Mon Mar 24 11:18:42 1997
 Job time : 11 secs.

US-08-644-289-1 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Mar 24 11:18:02 1997; MacPer time 2.66 Seconds

193,564 Million cell updates/sec

Tabular output not generated.

Title: >US-08-644-289-1

Description: (1-20) from US08644289.pep

Perfect Score: 151

Sequence: 1 SARPEKALVNEKGHRPSHC 20

Scoring table:

PAM 150

Gap 15

Searched: 82182 segs, 25727515 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

pir48
1:am1 2:am2 3:am3 4:unam1 5:unam2 6:unam3 7:unam4
8:unam5 9:unam6 10:unam7 11:unam8 12:unam9 13:unamc
14:unrev

Statistics: Mean 27.793; Variance 45.751; scale 0.607

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	ID	Description	Pred. No.
1	65	43.0		245	14	S54125	PVIII protein - Avia	1.44e+00
2	64	42.4		116	8	A39432	ATP-dependent exonic	2.08e+00
3	62	41.1		164	12	S14016	X16 protein - mouse	4.32e+00
4	62	41.1		1873	12	A55645	voltage-dependent ca	4.32e+00
5	60	39.7		116	5	E30563	T-cell receptor beta	8.83e+00
6	60	39.7		247	1	PRMSCL	cytotoxic T-lymphocy	8.83e+00
7	60	39.7		347	12	S12955	calcium channel prot	8.83e+00
8	60	39.7		466	9	A54125	ferrochelatase (EC 4	8.83e+00
9	60	39.7		1331	12	S05011	calcium channel alph	8.83e+00
10	60	39.7		1610	12	A46227	voltage-dependent Ca	8.83e+00
11	60	39.7		1646	12	JH0422	voltage-dependent ca	8.83e+00
12	60	39.7		1687	14	S41742	calcium channel alph	8.83e+00

13	60	39.7	1873	12	A30063	dihydropyridine rece	8.83e+00
14	60	39.7	2139	12	A44467	voltage-dependent ca	8.83e+00
15	60	39.7	2140	12	JH0426	voltage-dependent ca	8.83e+00
16	60	39.7	2143	12	JH0427	voltage-dependent ca	8.83e+00
17	60	39.7	2161	11	JH0564	calcium channel alph	8.83e+00
18	60	39.7	2166	14	S11339	calcium channel prot	8.83e+00
19	60	39.7	2171	12	S05054	calcium channel alph	8.83e+00
20	60	39.7	2181	11	A38198	calcium channel alph	8.83e+00
21	60	39.7	2208	11	A37860	calcium channel prot	8.83e+00
22	60	39.7	2220	11	A45290	calcium channel prot	8.83e+00
23	59	39.1	299	10	S44554	probable membrane pr	1.26e+01
24	59	39.1	299	9	A55890	citrate transport pr	1.26e+01
25	59	39.1	392	7	S15295	nolc protein - Rhizo	1.26e+01
26	58	38.4	110	5	A30519	T-cell receptor beta	1.78e+01
27	58	38.4	115	5	A36296	T-cell receptor beta	1.78e+01
28	58	38.4	124	9	A41139	protein kinase 1 (EC	1.78e+01
29	58	38.4	131	5	A27022	T-cell receptor beta	1.78e+01
30	58	38.4	416	9	A45510	probable protein kin	1.78e+01
31	58	38.4	444	11	C55886	dopamine receptor D1	1.78e+01
32	58	38.4	480	8	B61213	hypothetical protein	1.78e+01
33	58	38.4	531	9	B30311	protein kinase C (EC	1.78e+01
34	58	38.4	578	9	JC1385	protein kinase (EC 2	1.78e+01
35	58	38.4	586	9	JN0505	protein kinase C (EC	1.78e+01
36	58	38.4	609	9	A30311	protein kinase C (EC	1.78e+01
37	58	38.4	1157	3	GMLJIK	pol polyprotein - si	1.78e+01
38	57	37.7	193	3	ASLJST	vif protein - simian	2.51e+01
39	57	37.7	336	1	DEECDO	dihydroxycate oxida	2.51e+01
40	57	37.7	410	11	A44391	serum response elme	2.51e+01
41	57	37.7	502	12	A55197	Hiskot-Altrich synd	2.51e+01
42	57	37.7	604	10	S15794	lin-9 protein - Caen	2.51e+01
43	56	37.1	922	6	A30816	band 3 anion transp	2.51e+01
44	56	37.1	487	3	VZE8PT	virulence membrane p	3.52e+01
45	56	37.1	624	12	S41688	DNA-binding protein	3.52e+01

ALIGNMENTS

RESULT 1
ENTRY S54125 #type complete
TITLE PVIII protein - Avian adenovirus
ORGANISM #formal name Avian adenovirus
DATE 08-Jul-1995 #sequence_revision 08-Jul-1995 #text_change 08-Jul-1995

ACCESSIONS S54125
REFERENCE S54125
#authors Hess, M.; Cuzange, A.; Chroboczek, J.; Ruitrok, R.; Jactot, B.

#submission submitted to the EMBL Data Library, February 1995
#description The sequence of the two fibers of an avian adenovirus (CELO) and organization of the genome.

#accession S54125
#status preliminary
#residues 1-245 #label HES
#cross-references EMBL:X84724
SUMMARY #length 245 #molecular-weight 26877 #checksum 1107

Query Match 43.0%; Score 65; DB 14; Length 245;
Best Local Similarity 40.0%; Pred. No. 1.44e+00;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

DB 145 rplvgrstrprwc 159
QY 6 KALVNEKGHRPSHC 20

```

RESULT      2
ENTRY
TITLE      A39432      #type complete
            ATP-dependent exonuclease synthesis protein AddB - Bacillus
            subtilis
ORGANISM    #formal_name Bacillus subtilis
DATE        21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change
            23-Mar-1993
ACCESSIONS  A39432
REFERENCE   A39432
            Koistira, J.; Venema, G.
            J. Bacteriol. (1991) 173:3644-3655
            Cloning, sequencing, and expression of Bacillus subtilis
            genes involved in ATP-dependent nuclease synthesis.
            #cross-references MUID:91267926
            #accession A39432
            #status preliminary
            #molecule_type DNA
            #residues 1-1166 #label K00
            #cross-references GB:M33489
SUMMARY     #length 1166 #molecular-weight 134630 #checksum 5494

Query Match      42.4%; Score 64; DB 8; Length 1166;
Best Local Similarity 69.2%; Pred. No. 2.08e+00;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db      329 reihaltvrekgyr 341
      1: ||||| |
      3 RPPKALVREKGR 15

```

RESULT	3	
ENTRY	S14016	#type complete
TITLE	X16 protein - mouse	
ORGANISM	#formal_name Mus musculus	#common_name house mouse
DATE	18-Feb-1994	#sequence_revision 10-Nov-1995
	17-Nov-1995	#text_change
ACCESSIONS	S14016	
REFERENCE	S14016	
#authors	Ayane, M.; Preuss, U.; Koehler, G.; Nielsen, P.J.	
#journal	Nucleic Acids Res. (1991) 19:1273-1278	
#title	A differentially expressed murine RNA encoding a protein with similarities to two types of nucleic acid binding motifs.	
#cross-references	EMBL:91232908	
#accession	S14016	
#molecule_type	mRNA	
#residues	1-164	#label AY1
#cross-references	EMBL:X53824	
CLASSIFICATION	#superfamily ribonucleoprotein repeat	homology
FEATURE		
1-73	#domain ribonucleoprotein repeat	homology
SUMMARY	#length 164	#molecular_weight 19329
		#checksum 8302
Query Match	41.1%;	Score 62; DB 12; Length 164;
Best Local Similarity	50.0%;	Pred. No. 4.3e+00;
Matches	7;	Conservative
	5;	Mismatches
	2;	Indels
	0;	Gaps
0;		
Dc	137	rslstcrnhkpsars 150
	::: :: :: ::	
Qy	6	KALVREKGRSHS 19
RESULT	4	
ENTRY	A55645	#type complete

```

TITLE          voltage-dependent calcium channel alpha-1 chain - human
ORGANISM      #formal_name Homo sapiens #common_name man
DATE          24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change
              01-Mar-1996

ACCESSIONS    A55645
REFERENCE      A55645
               Hogan, K.; Powers, P.A.; Gregg, R.G.
               Genomics (1994) 24:608-609
               Cloning of the human skeletal muscle alpha-1 subunit of the
               dihydropyridine-sensitive L-type calcium channel
               (CACNL1A3) .

#accession    A55645
#status       preliminary
#molecule_type mRNA
#residues     1-1873 #label HOG
#cross-references GB:L33798

GENETICS      CDB:CACNL1A3
               #cross-references CDB:G00-126-431
               #map_position 1q31-q32

SUMMARY       #length 1873 #molecular-weight 212301 #checksum 8406

Query Match          41.1%; Score 62; DB 12; Length 1873;
Best Local Similarity 42.1%; Pred. No. 4.3e+00;
Matches      8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db          902 Lrp1rainrakglhvarc 920
||||:|:| | | | |
          2 LRPKALVREKGRHPSHSC 20

```

ENTRY	5	E30563	#type fragment
TITLE		T-cell receptor beta chain V region (CRTB3) - rat (fragment)	
ORGANISM		#formal name Rattus norvegicus #common name Norway rat	
DATE		03-Apr-1989 #sequence_revision 03-Apr-1989 #text_change 12-Apr-1995	
ACCESSIONS		E30563	
REFERENCE		A30563	
#authors		Williams, C.B.; Gutman, G.A.	
#journal		J. Immunol. (1989) 142:1027-1035	
#title		T cell receptor beta-chain genes in the rat. Availability and pattern of utilization of V gene segments differs from that in the mouse.	
#cross-references	MOID:89110038		
#accession	E30563		
#status		preliminary; not compared with conceptual translation	
#molecule_type	DNA		
#residues	1-116	#label	WIL
#note		the sequence was determined from the differentiated gene	
CLASSIFICATION		#superfamily immunoglobulin V region; immunoglobulin homology	
SUMMARY		length 116	#checksum 1551
Query Match		39.7%;	Score 60; DB 5; length 116;
Best Local Similarity		53.8%;	Pred. No. 8.83e+00;
Matches	7;	Conservative	4; Mismatches 2; Gaps 0;
DQ	12	Likeqgrvmsc	24
		:	
QY	8	LYRKGRPSHC	20
RESULT	6		
ENTRY		PRMSCL	#type complete

Mar 24 11:16

US-08-644-289-1.rpt

5

TITLE cytotoxic T-lymphocyte proteinase 1; cytotoxic T-cell-specific mouse

ALTERNATE_NAMES CPPI; cytotoxic cell protease 1; cytotoxic T-cell-specific protease 1

ORGANISM #formal name Mus musculus #common name house mouse

DATE 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 14-Jul-1994

ACCESSIONS A94288; A93382; A28952; A00956

REFERENCE A94288

authors Lobe, C.G.; Finlay, B.B.; Paranchych, W.; Paetkau, V.H.; Bleackley, R.C.

journal Science (1986) 232:858-861

title Novel serine proteases encoded by two cytotoxic T lymphocyte-specific genes.

cross-references M01D:86208120

accession A94288

cross-references M01D:86208120

residues 1-247 #label L08

cross-references GB:M12302

REFERENCE A93382

authors Brune, J.F.; Dosseto, M.; Denizot, F.; Mattei, M.G.; Clark, W.R.; Haggl, T.M.; Ferrier, P.; Nabholz, M.; Schmitt-Verhulst, A.M.; Luciani, M.F.; Golstein, P.

journal Nature (1986) 322:268-271

title The inducible cytotoxic T-lymphocyte-associated gene transcript CTLA-1 sequence and gene localization to mouse chromosome 14.

cross-references M01D:86284960

contents chromosome mapping

accession A93382

molecule_type mRNA

residues 1-247 #label BRU

cross-references EMBL:X04072

REFERENCE A90536

authors Lobe, C.G.; Upton, C.; Duggan, B.; Ehrman, N.; Letellier, M.; Bell, J.; McFadden, G.; Bleackley, R.C.

journal Biochemistry (1988) 27:6941-6946

title Organization of two genes encoding cytotoxic T lymphocyte-specific serine proteases CPPI and CPPII.

cross-references M01D:89062424

accession A28952

molecule_type DNA

residues 1-247 #label L02

cross-references GB:M2526

COMMENT This enzyme is probably necessary for target cell lysis in cell-mediated immune responses.

GENETICS

map_position 14

introns 19/1; 68/2; 113/3; 200/3

CLASSIFICATION #superfamily trypsin; trypsin homology

KEYWORDS hydrolase; serine proteinase; T-cell

FEATURE

1-18 #domain signal sequence #status predicted #label SIG1

19-20 #domain propeptide #status predicted #label APY

21-247 #product cytotoxic T-lymphocyte proteinase 1 #status predicted #label MPT

21-240 #domain trypsin homology #label TRY

49-65,142-209,173-188 #disulfide bonds #status predicted

64,108,203 #active site His, Asp, Ser #status predicted

SUMMARY #length 247 #molecular-weight 27470 #checksum 92

Query Match 39.7%; Score 60; DB 1; Length 247;

Best Local Similarity 38.9%; Pred. No. 8.83e+00;

Mar 24 11:16

US-08-644-289-1.rpt

6

Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db 32 rymallsikdqpaeic 49

Qy 3 RPFKALVREKGRHSNC 20

RESULT 7

ENTRY S12955 #type complete

TITLE calcium channel protein type L - rabbit

ORGANISM #formal name Oryctolagus cuniculus #common name domestic rabbit

DATE 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995

ACCESSIONS S12955

REFERENCE S12955

authors Huang, P.; Temizer, D.; Quetters, T.

journal FEBS Lett. (1990) 274:207-213

title Polymerase chain reaction cloning of L-type calcium channel sequences from the heart and the brain.

cross-references M01D:91071434

accession S12955

status preliminary

molecule_type mRNA

residues 1-347 #label H1A

cross-references EMBL:X58696

SUMMARY #length 347 #molecular-weight 40393 #checksum 7147

Query Match 39.7%; Score 60; DB 12; Length 347;

Best Local Similarity 42.1%; Pred. No. 8.83e+00;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 7 lrp1rainrakqlkhvqc 25

Qy 2 LRPFKALVREKGRHSNC 20

RESULT 8

ENTRY A54125 #type complete

TITLE ferrochelatase (EC 4.99.1.1) precursor, chloroplast - Arabidopsis thaliana

ORGANISM #formal name Arabidopsis thaliana #common name mouse-ear cress

DATE 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 07-Jul-1995

ACCESSIONS A54125

REFERENCE A54125

authors Smith, A.G.; Santana, M.A.; Wallace-Cook, A.D.M.; Roper, J.M.; Labbe-Bois, R.

journal J. Biol. Chem. (1994) 269:13405-13413

title Isolation of a cDNA encoding chloroplast ferrochelatase from Arabidopsis thaliana by functional complementation of a yeast mutant.

accession A54125

status preliminary

molecule_type mRNA

residues 1-466 #label SWI

cross-references GB:X73417

KEYWORDS chloroplast; lyase

SUMMARY #length 466 #molecular-weight 52032 #checksum 6060

Query Match 39.7%; Score 60; DB 9; Length 466;

Best Local Similarity 43.8%; Pred. No. 8.83e+00;

Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Mar 24 11:16

US-08-644-289-1.mpr

7

Db 10 lrp1lrairakqlkhvpsc 25
|:::| | | | |
Qy 5 lrpfkalvrekghrpspsc 20

RESULT 9
ENTRY S05011 #type fragment
TITLE calcium channel alpha-1 chain, dihydropyridine sensitive,
cardiac (clone HTDHP 2.0) - rabbit (fragment)
ALTERNATE_NAMES dihydropyridine receptor, cardiac; voltage-dependent calcium
channel type I
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
rabbit
DATE 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
18-Jun-1993

ACCESSIONS S05011
REFERENCE S05011
#authors Slish, D.F.; Engle, D.B.; Varadi, G.; Lotan, I.; Singer, D.;
Dascal, N.; Schwartz, A.
#journal FEBS Lett. (1989) 250:509-514
#title Evidence for the existence of a cardiac specific isoform of
the alpha(1) subunit of the voltage dependent calcium
channel.
#cross-references M01D:89325623
#accession S05011
#status not compared with conceptual translation
#molecule_type mRNA
#residues 1-1331 #label SLI
SUMMARY #length 1331 #checksum 8010

Query Match 39.7%; Score 60; DB 12; Length 1331;
Best Local Similarity 42.1%; Pred. No. 8.83e+00;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 293 lrp1lrairakqlkhvpsc 311
|:::| | | | |
Qy 2 lrpfkalvrekghrpspsc 20

RESULT 10
ENTRY A46227 #type complete
TITLE voltage-dependent Ca2+ channel alpha 1-subunit - golden
hamster
ORGANISM #formal_name Mesocricetus auratus #common_name golden hamster
DATE 27-Oct-1993 #sequence_revision 18-Nov-1994 #text_change
05-Apr-1995

ACCESSIONS A46227
REFERENCE A46227
#authors Yanev, G.C.; Wheeler, M.B.; Wei, X.; Perez-Reyes, E.;
Birnbauer, L.; Boyd III, A.E.; Moss, L.G.
#journal Mol. Endocrinol. (1992) 6:2143-2152
#title Cloning of a novel alpha 1-subunit of the voltage-dependent
calcium channel from the beta-cell.
#cross-references M01D:93149124
#accession A46227
#status preliminary; not compared with conceptual translation
#molecule_type mRNA
#residues 1-1610 #label YAN
#cross-references NCBI:123692
#experimental_source insulin-secreting cell line HIT-T15
#note sequence extracted from NCBI backbone
SUMMARY #length 1610 #molecular-weight 182326 #checksum 806

Mar 24 11:16

US-08-644-289-1.mpr

8

Query Match 39.7%; Score 60; DB 12; Length 1610;
Best Local Similarity 42.1%; Pred. No. 8.83e+00;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 987 lrp1lrairakqlkhvpsc 1005
|:::| | | | |
Qy 2 lrpfkalvrekghrpspsc 20

RESULT 11
ENTRY JH0422 #type complete
TITLE voltage-dependent calcium channel complex alpha-1 chain - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change
08-Feb-1996

ACCESSIONS JH0422; D35901
REFERENCE JH0422
#authors Hui, A.; Ellinor, P.T.; Krizanova, O.; Wang, J.J.; Diebold,
R.J.; Schwartz, A.
#journal Neuron (1991) 7:35-44
#title Molecular cloning of multiple subtypes of a novel rat brain
isoform of the alpha 1 subunit of the voltage-dependent
calcium channel.
#cross-references M01D:91299338
#accession JH0422
#molecule_type mRNA
#residues 1-1646 #label HUI
#cross-references GB:M57682
#experimental_source brain

REFERENCE A35901
#authors Snutch, T.P.; Leonard, J.P.; Gilbert, M.M.; Lester, H.A.;
Davidson, N.
#journal Proc. Natl. Acad. Sci. U.S.A. (1990) 87:3391-3395
#title Rat brain expresses a heterogeneous family of calcium
channels.
#cross-references M01D:90239020
#accession D35901
#status preliminary; not compared with conceptual translation
#molecule_type mRNA
#residues 1247-1434 #label SNU
#experimental_source brain

COMMENT #note the nucleotide sequence is not given
Calcium channels are essential for many cellular functions, such as
muscle contraction, propagation of action potentials, maintenance
of electrical activity, and neurotransmitter regulation.
KEYWORDS glycoprotein
FEATURE
1463-1491 #domain calcium-binding #status predicted #label EFC\
154,224,328 #binding_site carbohydrate (Asn) (covalent) #status
predicted\
464,848,1489,1584 #binding_site phosphate (Ser) (covalent) #status
predicted

SUMMARY #length 1646 #molecular-weight 186842 #checksum 699

Query Match 39.7%; Score 60; DB 12; Length 1646;
Best Local Similarity 42.1%; Pred. No. 8.83e+00;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 988 lrp1lrairakqlkhvpsc 1006
|:::| | | | |
Qy 2 lrpfkalvrekghrpspsc 20

RESULT 12

Mar 24 11:16

US-08-644-289-1.rpt

9

```

ENTRY      541742  #type complete
TITLE      calcium channel alpha-1-chain - House fly
ORGANISM   #formal name Musca domestica #common name house fly
DATE       25-Dec-1994 #sequence_revision 25-Dec-1994 #text_change
          25-Dec-1994
ACCESSIONS 541742
REFERENCE  541742
          Grabner, M.; Bachmann, A.; Rosenthal, F.; Strlesnig, J.;
          Schulz, C.; Tautz, D.; Glosemann, H.
          FEBS Lett. (1994) 339:189-194
          #journal
          #title      Insect calcium channels. Molecular cloning of an alpha
          (1)-subunit from housefly (Musca domestica) muscle.
          #accession  541742
          #status     preliminary
          #residues   1-1687 #label GRA
SUMMARY    #length 1687 #molecular-weight 193873 #checksum 9300

Query Match      39.7%; Score 60; DB 14; Length 1687;
Best Local Similarity 42.1%; Pred. No. 8.83e+00;
Matches          8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db      893 lrp1rainrakqlkhvqc 911
      |||::|||:| | | |
      2 LRPFKALVREKGRPSHSC 20

RESULT 13
ENTRY   A30063  #type complete
TITLE   dh5alpha deoxyribonuclease III - rabbit
ORGANISM #formal name Oryctolagus cuniculus #common name domestic
          rabbit
DATE    30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change
          18-Jun-1993
ACCESSIONS A30063
REFERENCE  A30063
          Tanabe, T.; Takeshima, H.; Mikami, A.; Flockerzi, V.;
          Takahashi, H.; Kangawa, K.; Kojima, M.; Matsuo, H.; Hirose,
          T.; Numa, S.
          #journal
          #title      Nature (1987) 328:313-318
          Primary structure of the receptor for calcium channel
          blockers from skeletal muscle.
          #cross-references M01D:87258269
          #accession  A30063
          #molecule_type mRNA
          #residues   1-1873 #label TAN
SUMMARY    #length 1873 #molecular-weight 212027 #checksum 5776

Query Match      39.7%; Score 60; DB 12; Length 1873;
Best Local Similarity 42.1%; Pred. No. 8.83e+00;
Matches          8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db      902 lrp1rainrakqlkhvqc 920
      |||::|||:| | | |
      2 LRPFKALVREKGRPSHSC 20

RESULT 14
ENTRY   A44467  #type complete
TITLE   voltage-dependent calcium channel complex alpha-1 chain mbc -
          mouse
ORGANISM #formal name Mus musculus #common name house mouse
DATE     31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
          03-Mar-1995
ACCESSIONS A44467

```

Mar 24 11:16

US-08-644-289-1.rpt

10

```

REFERENCE  A44467
          Ma, W.J.; Holz, R.W.; Uhler, M.D.
          #authors
          #journal    J. Biol. Chem. (1992) 267:22728-22732
          #title      Expression of a cDNA for a neuronal calcium channel alpha 1
          subunit enhances secretion from adrenal chromaffin cells.
          #cross-references M01D:93054582
          #accession  A44467
          #molecule_type mRNA
          #residues   1-2139 #label MA1
          #cross-references NCBI:118042
          #experimental_source brain
          #note        sequence extracted from NCBI backbone
KEYWORDS   glycoprotein
SUMMARY    #length 2139 #molecular-weight 240136 #checksum 8452

Query Match      39.7%; Score 60; DB 12; Length 2139;
Best Local Similarity 42.1%; Pred. No. 8.83e+00;
Matches          8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db      1003 lrp1rainrakqlkhvqc 1021
      |||::|||:| | | |
      2 LRPFKALVREKGRPSHSC 20

RESULT 15
ENTRY   JH0426  #type complete
TITLE   voltage-dependent calcium channel complex alpha-1 chain rbc-1
          - rat
ORGANISM #formal name Rattus norvegicus #common name Norway rat
DATE     31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change
          08-Feb-1996
ACCESSIONS JH0426; C35901
REFERENCE  JH0426
          Snutch, T.P.; Tomlinson, W.J.; Leonard, J.P.; Gilbert, M.M.
          #authors
          #journal    Neuron (1991) 7:45-57
          #title      Distinct calcium channels are generated by alternative
          splicing and are differentially expressed in the mammalian
          CNS.
          #cross-references M01D:91299339
          #accession  JH0426
          #molecule_type mRNA
          #residues   1-2140 #label SNU
          #cross-references GB:M57516
          #experimental_source brain
          #note        the codon for 1758-Tyr is TAA; the authors claim that
          this is an artifact of the cDNA cloning. Another cDNA
          clone contains TAC at this position
REFERENCE  A35901
          Snutch, T.P.; Leonard, J.P.; Gilbert, M.M.; Lester, H.A.;
          Davidson, N.
          #journal    Proc. Natl. Acad. Sci. U.S.A. (1990) 87:3391-3395
          #title      Rat brain expresses a heterogeneous family of calcium
          channels.
          #cross-references M01D:90239020
          #accession  C35901
          #status     preliminary; not compared with conceptual translation
          #molecule_type mRNA
          #residues   1139-1384 #label SN2
          #note        the nucleotide sequence is not given
COMMENT    Calcium channels are essential for many cellular functions, such as
          muscle contraction, propagation of action potentials, maintenance
          of electrical activity, and neurotransmitter regulation.
FEATURE    1545,1597,1670,

```

Mar 24 11:16

US-08-644-289-1.rpt

11

1818,1898

#binding_site phosphate (Ser) (covalent) #status
predicted

SUMMARY

#length 2140 #molecular-weight 240188 #checksum 1204

Query Match

39.7%; Score 60; DB 12; Length 2140;

Best Local Similarity 42.1%; Pred. No. 8.83e+00;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 1003 lrp1rainrakg1khvqc 1021

||||:|:| | | : |

Oy 2 LRPFKALVREKCHRP SHSC 20

Search completed: Mon Mar 24 11:18:15 1997
Job time : 13 secs.

(a) (b) (c) (d) (e) (f)

(7M)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

```

MPerch_mn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Mar 25 02:41:07 1997; MasPar time 58.62 Seconds
Tabular output not generated. 533.919 Million cell updates/sec

```

Title:	>US-08-644-289-2
Description:	(1-38) from US08644289.seq
Perfect Score:	38
N.A. Sequence:	1 ACTCAGCCCTTAGAATTAAAGATGCCCATGTCACAGA 38
Comp:	TGAGTCGGGAATCTAATTTTCCTCAGCGTACATGCTCT

Scoring table: **TABLE default**
Gap 10

```
Nmatch  STD :  Dbase 0;  Query 0
```

Searched: 279077 seqs, 411808665 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: **embl-new5**

Database: genbank94

16:BCT1 1:1:BCT2 18:BCT3 19:BCT4 20:BCT5 21:BCT6 22:BCT7
23:BCT8 24:BCT9 25:INV1 26:INV2 27:INV3 28:INV4 29:INV5
30:INV6 31:INV7 32:NAM1 33:NAM2 34:NAM3 35:VR1 36:VR2
37:VAT3 38:PMT1 39:PMT2 40:PMT3 41:PHG 42:PLNT1 43:PLNT2
44:PLN3 45:PLN4 46:PLN5 47:PLN6 48:PLN7 49:PLN8 50:PRIL1
51:PR12 52:PR13 53:PR14 54:PR15 55:PR16 56:PR17 57:PR18
58:PR19 59:PR110 60:PR11 61:PR112 62:PR113 63:ROD1
64:ROD2 65:ROD3 66:ROD4 67:ROD5 68:ROD6 69:ROD7 70:ROD08
71:STR 72:SN 73:ONA 74:VAL1 75:VAL2 76:VAL3 77:VAL4
78:VAL5 79:VAL6 80:VAL7 81:VAL8

Database: genbank-new5

Database: u-emb146_94 .

Statistics: Mean 7.583; Variance 2.828; scale 2.681

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				ALIGNMENTS	
Result No.	Score	Query Match Length	ID	Description	Pred. No
1	34	89.5	107 63	WMANT10	4,26e-14
2	34	89.5	MUSP53B	Mouse p53 mRNA, comp1	4,26e-14
3	34	89.5	MUSP53C	Mouse p53 mRNA, compl	4,26e-14
4	34	89.5	MUSP53C	Mouse p53 mRNA, compl	4,26e-14
5	34	89.5	MUSP53M	Mouse p53 cellular tr	4,26e-14
6	34	89.5	MUSP53M	Mouse p53 cellular tr	4,26e-14
7	34	89.5	MUSP53Cg	Mouse p53 cellular tu	4,26e-14
8	34	89.5	MUSP53P	Mouse pseudogene for	4,26e-14
9	34	89.5	RATP53S07	Rattus norvegicus tum	2,51e-05
10	24	63.2	RNIO7019	Rattus norvegicus Wts	2,51e-05
11	24	63.2	RNP53	Rattus norvegicus Wts	2,51e-05
12	24	63.2	RNP53	Rattus norvegicus Wts	2,51e-05
13	20	52.6	RNP53	Rattus norvegicus Wts	3,66e-02
14	20	52.6	RNP53	Rattus norvegicus Wts	3,66e-02
15	20	52.6	RNP53	Rattus norvegicus Wts	3,66e-02
16	19	50.0	HUMP5310	Human cellular phosph	2,04e-01
17	19	50.0	HUMP5310	Human cellular phosph	2,04e-01
18	19	50.0	HSP53006	Human mRNA for mutate	2,04e-01
19	19	50.0	HSP53002	Human mRNA for mutate	2,04e-01
20	19	50.0	HSP53011	Human mRNA for mutate	2,04e-01
21	19	50.0	HSP53008	Human mRNA for mutate	2,04e-01
22	19	50.0	HSP53009	Human mRNA for mutate	2,04e-01
23	19	50.0	HSP53010	Human mRNA for mutate	2,04e-01
24	19	50.0	HSP53003	Human mRNA for mutate	2,04e-01
25	19	50.0	HSP53007	Human mRNA for mutate	2,04e-01
26	19	50.0	HSP53004	Human mRNA for mutate	2,04e-01
27	19	50.0	HSP53005	Human mRNA for mutate	2,04e-01
28	19	50.0	HUMP53B	Human p53 cellular tr	2,04e-01
29	19	50.0	HUMP53A	Human p53 cellular tr	2,04e-01
30	19	50.0	HSP53	Human mRNA for p53 ce	2,04e-01
31	19	50.0	CAP53	African Green Monkey	2,04e-01
32	19	50.0	HUMP53T	Human p53 cellular tr	2,04e-01
33	19	50.0	HSP53R	Human mRNA fragment f	2,04e-01
34	19	50.0	FCP53	Cat Lymph node mRNA f	2,04e-01
35	19	50.0	CATP53	Cat Lymph node mRNA f	2,04e-01
36	19	50.0	CATP53	Cat Lymph node mRNA f	2,04e-01
37	19	50.0	MUSP53A	Rhesus monkey p53 mRN	2,04e-01
38	19	50.0	CASBREML	CAS-BR-E mutine leuko	1,08e+00
39	18	47.4	CATP53A	Cat mRNA for p53 prot	5,41e+00
40	17	44.7	PHIX17A	P. phloerophagen DNA for	5,41e+00
41	17	44.7	PHIX17A	Bacteriophage phi-X17	5,41e+00
42	17	44.7	PHIX17A	Bacteriophage phi-X17	5,41e+00
43	17	44.7	PHIX17A	Bacteriophage phi-X17	5,41e+00
44	17	44.7	RATYRC5	Rat s-myc protein gen	5,41e+00
45	17	44.7	MUSYRCICg	Mouse XRC01 DNA repla	5,41e+00

Mar 25 02:40

US-08-644-289-2.rpt

3

Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 107)
AUTHORS Bienz, B., Zakut-Houri, R., Givol, D. and Oren, M.
TITLE Analysis of the gene coding for the murine cellular tumour antigen p53
JOURNAL EMBO J. 3 (9), 2179-2183 (1984)
MEDLINE 85027173
REFERENCE 2 (bases 1 to 107)
AUTHORS Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Givol, D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306 (5943), 594-597 (1983)
MEDLINE 84068204
FEATURES
source Location/Qualifiers
1..107
/organism="Mus musculus"
exon 1..107
/label=ex10
/usedin=X00876:P53_CDS
/usedin=X00876:P53_mRNA
BASE COUNT 27 a 25 c 35 g 20 t
ORIGIN

Query Match 89.5%; Score 34; DB 63; Length 107;
Best Local Similarity 100.0%; Pred. No. 4,28e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 44 aggccttagagtaagatgccatgctacaga 77
|||||
Qy 5 AGCCTTAGAGTTAAAGATGCCATGCTACAGA 38

RESULT 2
LOCUS MUSP53B 1241 bp mRNA ROD 02-NOV-1992
DEFINITION Mouse p53 mRNA, complete cds, clone p53-m1.
ACCESSION M13873
KEYWORDS g200200
SOURCE p53 gene.
ORGANISM Mus musculus
Mus (BALB/c) Meth A library, cDNA to mRNA, clone p53-m1.
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE 1 (bases 1 to 1241)
AUTHORS Arai, N., Nomura, D., Yokota, K., Wolf, D., Brill, E., Shohat, O. and Rotter, V.
TITLE Immunologically distinct p53 molecules generated by alternative splicing
JOURNAL Mol. Cell. Biol. 6, 3232-3239 (1986)
MEDLINE 87064640
FEATURES
source Location/Qualifiers
1..1241
/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/sequenced_mol="cDNA to mRNA"
/tissue_lib="Meth A"
69..1241
/gene="p53"
/map="11"
/codon_start=1
/db_xref="PID:g200201"
/translation="MTAMEESQSDISLEPLISQETFGMLKLPPEDILPSPHCMIDL
ILPDVVEFFEGSRALRVSGAPADPVERGVAAPATPAPLSSFPVSGKTYGG
NYGHGFLQSGTAKSVWCTYSPILKFLCQLAKTCVQVLNVSATPPAGSRVAMATY

Mar 25 02:40

US-08-644-289-2.rpt

4

KSQHMTGVRCPEHHERCSGDGLAPQHLIRVEGNLYPEYLEDQTRHVVPE
PPRAGEYTTIHYKWNSSCGAMRRPILTTITLESNGLLGDSFEVRCVACPG
RDRRETFPRKKEVLCPELPPSGARALPCTTSASPPQKKPLDGEYETLIKIRGR
FEMREINALELKDAMHTESGDSRAHSYLTAKKGSTSRHKTKMKVGPDSQ

BASE COUNT 289 a 316 c 321 g 255 t
ORIGIN

Query Match 89.5%; Score 34; DB 67; Length 1241;
Best Local Similarity 100.0%; Pred. No. 4,28e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1096 aggccttagagtaagatgccatgctacaga 1129
|||||
Qy 5 AGCCTTAGAGTTAAAGATGCCATGCTACAGA 38

RESULT 3
LOCUS MUSP53A 1285 bp mRNA ROD 05-NOV-1992
DEFINITION Mouse p53 mRNA, complete cds, clone pcd53.
ACCESSION M13872
KEYWORDS g200198
SOURCE p53 gene.
ORGANISM Mus musculus
Mus (BALB/c) nontransformed helper T-cell, cDNA to mRNA, clone pCD-p53.
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE 1 (bases 1 to 1285)
AUTHORS Arai, N., Nomura, D., Yokota, K., Wolf, D., Brill, E., Shohat, O. and Rotter, V.
TITLE Immunologically distinct p53 molecules generated by alternative splicing
JOURNAL Mol. Cell. Biol. 6, 3232-3239 (1986)
MEDLINE 87064640
FEATURES
source Location/Qualifiers
1..1285
/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/cell_type="nontransformed helper T-cell"
/sequenced_mol="cDNA to mRNA"
113..1285
/codon_start=1
/product="p53 protein"
/db_xref="PID:g200199"
/translation="MTAMEESQSDISLEPLISQETFGMLKLPPEDILPSPHCMIDL
ILPDVVEFFEGSRALRVSGAPADPVTETPPAPAPATPAPLSSFPVSGKTYGG
NYGHGFLQSGTAKSVWCTYSPILKFLCQLAKTCVQVLNVSATPPAGSRVAMATY
KKSQHMTGVRCPEHHERCSGDGLAPQHLIRVEGNLYPEYLEDQTRHVVPE
PPRAGEYTTIHYKWNSSCGAMRRPILTTITLESNGLLGDSFEVRCVACPG
RDRRETFPRKKEVLCPELPPSGARALPCTTSASPPQKKPLDGEYETLIKIRGR
FEMREINALELKDAMHTESGDSRAHSYLTAKKGSTSRHKTKMKVGPDSQ

Query Match 89.5%; Score 34; DB 67; Length 1285;
Best Local Similarity 100.0%; Pred. No. 4,28e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1140 aggccttagagtaagatgccatgctacaga 1173
|||||
Qy 5 AGCCTTAGAGTTAAAGATGCCATGCTACAGA 38

RESULT 4
LOCUS MUSEP53C 1322 bp mRNA ROD 02-NOV-1992
DEFINITION Mouse p53 mRNA, complete cde, clone p53-m8.
ACCESSION M13874
NID 9200202
KEYWORDS p53 gene.
SOURCE Mouse Ab-MuLV transformed Meth A fibroblast and normal T-cell line L12, cDNA to mRNA, clone p53-m8.
ORGANISM Mus musculus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE 1 (bases 1 to 1322)
AUTHORS Arai,N., Nomura,D., Yokota,K., Wolf,D., Brill,E., Shohat,O. and Rotter,V.
TITLE Immunologically distinct p53 molecules generated by alternative splicing
JOURNAL Mol. Cell. Biol. 6, 3232-3239 (1986)
MEDLINE 87064640
REFERENCE 2 (bases 1 to 1322)
AUTHORS Han,K.A. and Kulez-Martin,M.F.
TITLE Alternatively spliced p53 RNA in transformed and normal cells of different tissue types
JOURNAL Nucleic Acids Res. 20, 1979-1991 (1992)
MEDLINE 92253421
FEATURES
source
1..1322
/organism="Mus musculus"
/cell_line="normal T-cell line L12"
/cell_type="Ab-MuLV transformed Meth A fibroblast"
/sequenced_mol="cDNA to mRNA"
54..1199
/gene="p53"
/map="11"
/codon_start=1
/db_xref="PID:g200203"
/translation="MTAMEESQSDISLEPLISOETFGSLKLLPEDILPSPHCMDLLPRLDVEEFEECPSEALRVSQAPADPVTETRGVAPAPAPMPLSSFPQSKTYGNYGHLGFLQSGTAKSYMCTYSPPLNKLFLQIAKTCVQIWMASAPAGSRRAALYKSOHMTVEVRRCCHERCSDGDLAPQHLIRVEGNLIPYLEDROTFRHSVVPTEPPEAGSEYTTIHYKMCNSSCGMGNRRPILITITLEDSSGNLGRSFEVAVCAFCGRDRRTTEENFRKKEVLCPELPPGSAKRALPTCTSSPQKKRP LDGCEFTLIKRRRFEMRELNLEALDLDAATEESGDSRAHSLQPRAFQAIKEESPNC"
BASE COUNT 308 a 407 c 329 g 278 t
ORIGIN
Query Match 89.5%; Score 34; DB 67; Length 1322;
Best Local Similarity 100.0%; Pred. No. 4,28e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1081 aggccttagagttaaagatgcccatgtctacaga 1114
|||||
Qy 5 AGGCCTTAGAGTTAAAGATGCCCATGTCTACAGA 38
RESULT 5
LOCUS MUP53 1377 bp RNA ROD 12-SEP-1993
DEFINITION Mouse mRNA for transformation associated protein p53.
ACCESSION X00741
NID 953570
KEYWORDS oncogene.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; mitochondrial eukaryotes; Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;

REFERENCE
AUTHORS Jenkins,J.R., Rudge,K., Redmond,S. and Wade-Evans,A.
TITLE Cloning and expression analysis of full length mouse cDNA sequences encoding the transformation associated protein p53
JOURNAL Nucleic Acids Res. 12 (14), 5609-5626 (1984)
MEDLINE 84272240
COMMENT Data kindly reviewed (19-FEB-1986) by A. Wade-Evans.
FEATURES
source
1..1377
/organism="Mus musculus"
123..1292
/note="p53"
/codon_start=1
/db_xref="PID:g53571"
/translation="MTAMEESQSDISLEPLISOETFGSLKLLPEDILPSPHCMDLLPRLDVEEFEECPSEALRVSQAPADPVTETRGVAPAPAPMPLSSFPQSKTYGNYGHLGFLQSGTAKSYMCTYSPPLNKLFLQIAKTCVQIWMASAPAGSRRAALYKSOHMTVEVRRCCHERCSDGDLAPQHLIRVEGNLIPYLEDROTFRHSVVPTEPPEAGSEYTTIHYKMCNSSCGMGNRRPILITITLEDSSGNLGRSFEVAVCAFCGRDRRTTEENFRKKEVLCPELPPGSAKRALPTCTSSPQKKRP LDGCEFTLIKRRRFEMRELNLEALDLDAATEESGDSRAHSLYTKKQSTSNHKKTMKVKVPDSD"
BASE COUNT 316 a 422 c 348 g 291 t
ORIGIN
Query Match 89.5%; Score 34; DB 64; Length 1377;
Best Local Similarity 100.0%; Pred. No. 4,28e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1147 aggccttagagttaaagatgcccatgtctacaga 1180
|||||
Qy 5 AGGCCTTAGAGTTAAAGATGCCCATGTCTACAGA 38
RESULT 6
LOCUS MUSEP53M 1772 bp mRNA ROD 03-MAY-1985
DEFINITION Mouse p53 cellular tumor antigen, mRNA.
ACCESSION K01700
NID g200204
KEYWORDS DNA-binding protein; antigen; p53 gene; tumor antigen.
SOURCE Mouse embryo P9 carcinoma cells, cDNA to mRNA, clones p176, p271, p422, p208 [1], and clone pp53-1 [2].
ORGANISM Mus musculus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE 1 (bases 1 to 1716)
AUTHORS Zakut-Houri,R., Oren,M., Biern,B., Lavie,V., Hazum,S. and Gliori,D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306, 594-597 (1983)
MEDLINE 84068204
REFERENCE 2 (bases 105 to 1772)
AUTHORS Pennica,D., Goeddel,D.V., Hayflick,J.S., Reich,N.C., Anderson,C.W. and Levine,A.J.
TITLE The amino acid sequence of murine p53 determined from a cDNA clone
JOURNAL Virology 134, 477-482 (1984)
MEDLINE 86072076
COMMENT The murine p53 protein is similar to the avian and human myc gene products and the adenovirus E1a proteins. All three of these proteins are rich in proline, contain proline runs or clusters, and are localized in the nucleus of transformed cells. The E1a proteins and p53 have been shown to have very short half-lives (2). The sequence of cDNA in [1] was established by analysing four separate

cDNA clones; p176 (bp 1-1247), p271 (bp 69-638), p422 (bp 1288-1719), and p208 (bp 1432-1719) [1]. There is only one functional p53 gene in the mouse genome. All existing different forms of murine p53 must be products of the same gene, mostly due to post-translational modifications [1]. The sequence of the murine p53 pseudogene, also reported by [1] (see separate entry), and the cDNA sequence are almost identical from nucleotide 186 onward. Upstream of this position the two sequences diverge totally and no homology can be observed; downstream of here the two sequences differ by only 4%. The differences are due to substitutions and to some small deletions or additions in this gene relative to the cDNA [1].

FEATURES

source

1..1772
Location/Qualifiers
/organism="Mus musculus"
/cell_line="F9"
/dev_stage="embryo"
/sequenced_mol="cDNA to mRNA"
/tissue_type="carcinoma"
<1..1668
/gene="p53"
/map="11"
158..1330
/gene="p53"
/map="11"
/codon_start=1
/product="cellular tumor antigen"
/db_xref="PID:q20205"
/translation="MTAMEESQSDISLELPISQETFGSLMKLPPEDILPSPHCMDDL
LLPDVVEFEFGSEALRVSGAPADPVETEGVAPAPATPWPPLSSFPSSOKTYG
NYGFHLGFLGCTAKSYNCTYSPPLAKLEFCQLAKTCQVQLWVSAITPPASGRAMAIY
KKSQMTVEVARCPNHERCSDGDLAPQHLIRVEGNLYPELYEDQTRHSVVPYE
PPAGSEYTTIHYKYNCSGCMGNRRPILITITLEDSSGILGRDSEVRVACG
RDRRTDEENFRKEVLCPELPFGSAKRALPTCSAPPOKKRPIDGTYFTLAKRGR
FEMRELMEALELKDHAATEESGDSRAHSSTYLRKQSTSRHKTMVKVGPDS"

CDS

BASE COUNT 385 a 546 c 412 g 429 t
ORIGIN 165 bp upstream of NcoI site.

Query Match 89.5%; Score 34; DB 67; Length 1772;
Best Local Similarity 100.0%; Pred. No. 4.28e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1185 aggccttagagtaagatgccatgctacaga 1218
|||||
Qy 5 AGGCCTTAGATTAAAGATGCCATGCTACAGA 38

RESULT 7
LOCUS MPE53R 1773 bp RNA ROD 12-SEP-1993
DEFINITION Mouse mRNA for cellular tumour antigen p53.
ACCESSION X01231 K01700
ID 953575
KEYWORDS antigen; tumor antigen.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Zakut-Houri, R., Oren, M., Blenz, B., Lavie, V., Hazum, S. and Givol, D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306 (5943), 594-597 (1983)

MEDLINE 84068204
REFERENCE 2 (bases 1 to 1773)
AUTHORS Blenz, B., Zakut-Houri, R., Givol, D. and Oren, M.
TITLE Analysis of the gene coding for the murine cellular tumour antigen p53

JOURNAL EMBO J. 3 (9), 2179-2183 (1984)
MEDLINE 85027173
REFERENCE 3 (bases 1 to 1773)
AUTHORS Givol, R.

JOURNAL Direct Submission
JOURNAL Submitted (28-NOV-1985) to the EMBL/GenBank/DBJ databases
FEATURES
source
1..1773
Location/Qualifiers

1..1772
/organism="Mus musculus"
158..1330
/note="p53 polypeptide (aa 1-390)"
/codon_start=1
/db_xref="PID:q53576"

/translation="MTAMEESQSDISLELPISQETFGSLMKLPPEDILPSPHCMDDL
LLPDVVEFEFGSEALRVSGAPADPVETEGVAPAPATPWPPLSSFPSSOKTYG
NYGFHLGFLGCTAKSYNCTYSPPLAKLEFCQLAKTCQVQLWVSAITPPASGRAMAIY
KKSQMTVEVARCPNHERCSDGDLAPQHLIRVEGNLYPELYEDQTRHSVVPYE
PPAGSEYTTIHYKYNCSGCMGNRRPILITITLEDSSGILGRDSEVRVACG
RDRRTDEENFRKEVLCPELPFGSAKRALPTCSAPPOKKRPIDGTYFTLAKRGR
FEMRELMEALELKDHAATEESGDSRAHSSTYLRKQSTSRHKTMVKVGPDS"

old_sequence
/note="G is A in [1]"
/citation=[1]
404
/note="U is C in [1]"
/citation=[1]

old_sequence
/note="U is C in [1]"
/citation=[1]
561
/note="U is C in [1]"
/citation=[1]

old_sequence
/note="U is C in [1]"
/citation=[1]
635
/note="U is C in [1]"
/citation=[1]

old_sequence
/note="A is G in [1]"
/citation=[1]
660
/note="A is G in [1]"
/citation=[1]

old_sequence
/note="G is U in [1]"
/citation=[1]
859
/note="G is U in [1]"
/citation=[1]

old_sequence
/note="UA is CU in [1]"
/citation=[1]
1445..1446
/note="UA is CU in [1]"
/citation=[1]

old_sequence
/note="CU is UC in [1]"
/citation=[1]
1447..1448
/note="CU is UC in [1]"
/citation=[1]

old_sequence
/note="U is A in [1]"
/citation=[1]
1503
/note="U is A in [1]"
/citation=[1]

old_sequence
/note="CU is UC in [1]"
/citation=[1]
1558..1559
/note="CU is UC in [1]"
/citation=[1]

BASE COUNT 386 a 545 c 412 g 430 t
ORIGIN

Query Match 89.5%; Score 34; DB 64; Length 1773;
Best Local Similarity 100.0%; Pred. No. 4.28e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1185 aggccttagagtaagatgccatgctacaga 1218
|||||

Mar 25 02:40

US-08-644-289-2.rge

9

QY 5 AGCCTTAGAGTTAAGATGCCATGCTACAGA 38

RESULT 8 MUSP53PG 2130 bp DNA ROD 03-MAY-1985

LOCUS Mouse p53 cellular tumour antigen pseudogene.
DEFINITION K02110

ACCESSION g200206

KEYWORDS antigen; p53 gene; processed pseudogene; tumor antigen.
SOURCE Mouse 3.3-kb fragment isolated from a BALB/c genomic library, clone pch53-11.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2130)
AUTHORS Zakut-Houri,R., Oren,M., Bienz,B., Lavie,V., Hazum,S. and Givol,D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306, 594-597 (1983)MEDLINE 84068204
COMMENT The DNA sequence of pch53-11 contains a long poly-A tract, lacks introns, and is bounded by direct repeats (bp 169-181 and bp 1852-1864), suggesting that it is a processed gene which resulted from reverse transcription of the mature mRNA.

The sequence of the murine p53 cDNA, also reported by [1] (see separate entry), and the pseudogene are almost identical from nucleotide 186 onward. Upstream of this position the two sequences diverge totally and no homology can be observed; downstream of here the two sequences differ by only 4%. The differences are due to substitutions and to some small deletions or additions in this gene relative to the cDNA.

FEATURES Location/Qualifiers

source 1..2130

mRNA /organism="Mus musculus"
/note="pseudo-p53 mRNA"CDS 262..>1434
/note="pseudo-p53"
/pseudo
/codon_start=1BASE COUNT 602 a 585 c 456 g 487 t
ORIGIN 19 bp upstream of BglIII site.Query Match 89.5%; Score 34; DB 67; Length 2130;
Best Local Similarity 100.0%; Pred. No. 4.28e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;Db 1287 aggccttagagttaaagatgccatgctacaga 1320
|||||
QY 5 AGCCTTAGAGTTAAGATGCCATGCTACAGA 38RESULT 9 MHP53P 2132 bp DNA ROD 07-NOV-1985
LOCUS Mouse pseudogene for cellular tumour antigen p53.
DEFINITION X01236 K02110
ACCESSION g53574
NID antigen; pseudogene; tumor antigen.
KEYWORDS house mouse.
SOURCE Mus musculus

Eukaryotes; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Clitres; Rodentia;
Sciurognathi; Myomorpha; Muridae; Murinae; Mus.

Mar 25 02:40

US-08-644-289-2.rge

10

REFERENCE 1 (bases 1 to 2132)

AUTHORS Zakut-Houri,R., Oren,M., Bienz,B., Lavie,V., Hazum,S. and Givol,D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306 (5943), 594-597 (1983)

MEDLINE 84068204

FEATURES Location/Qualifiers

source 1..2132

misc_feature 1..185
/organism="Mus musculus"repeat_region 169..181
/note="5' flanking region without homology to p53 cDNA"misc_feature 186..1773
/note="direct repeat 1"misc_feature 1809..1814
/note="region of 96% homology to p53 cDNA"polyA_site 1828
/note="pot. polyadenylation signal"repeat_region 1854..1866
/note="direct repeat 1"BASE COUNT 602 a 586 c 457 g 487 t
ORIGINQuery Match 89.5%; Score 34; DB 64; Length 2132;
Best Local Similarity 100.0%; Pred. No. 4.28e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;Db 1288 aggccttagagttaaagatgccatgctacaga 1321
|||||

QY 5 AGCCTTAGAGTTAAGATGCCATGCTACAGA 38

RESULT 10

LOCUS RAFP53TS07 159 bp DNA ROD 14-JUL-1993

DEFINITION Rattus norvegicus tumor suppressor (p53) gene, exon 9.

ACCESSION 107909 L07781

NID g205949

KEYWORDS tumor suppressor.

SEGMENT 7 of 8

SOURCE Rattus norvegicus (strain Sprague-Dawley) DNA.

ORGANISM Rattus norvegicus

Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;

Eutheria; Rodentia; Myomorpha; Muridae; Murinae.

Hulla,U.E. and Schneider,R.P.Sr..

TITLE Structure of the rat p53 tumor suppressor gene

JOURNAL Nucleic Acids Res. 21, 713-717 (1993)

MEDLINE 93181268

FEATURES Location/Qualifiers

source 1..159

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/sequenced_mol="DNA"

order(L07908:315..340,1..26)

/gene="p53"

/note="intron 8 is ~830 bp"

/number=8

exon 27..133

/gene="p53"

/number=9

BASE COUNT 32 a 40 c 53 g 34 t
ORIGIN
Query Match 63.2%; Score 24; DB 69; Length 159;

Mar 25 02:40

US-08-644-289-2.rgc

11

Best Local Similarity 85.3%; Pred. No. 2.51e-05;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 70 aggccttggaatgaagatgccctgcgcaga 103
||||| || ||||| ||||| ||||| |||||
Oy 5 AGGCTTAGAGTTAAGAGTCCGCTACAGA 38

RESULT 11

LOCUS RNU07019 1273 bp DNA ROD 02-FEB-1996
DEFINITION Rattus norvegicus Wistar clone pR53p1 p53 pseudogene.
ACCESSION U07019
NID g460923

KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Myomorpha;
Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1273)
AUTHORS Lin, Y. and Chan, S.H.
TITLE Cloning and characterization of two processed p53 pseudogenes from the rat genome

JOURNAL Gene 156 (2), 183-189 (1995)

MEDLINE 95278743

REFERENCE 2 (bases 1 to 1273)

AUTHORS Lin, Y.
TITLE Direct Submision
SUBMITTED (22-FEB-1994) Yue Lin, Department of Microbiology,
National University of Singapore, Lower Kent Ridge Road, Singapore
0511, Singapore

FEATURES
source Location/Qualifiers

1..1273

/clone="pR53p1"

/strain="Wistar"

/organism="Rattus norvegicus"

/sex="male"

/cell_type="hepatocyte"

/tissue_type="liver"

<1..14

/note="corresponds to a part of exon 1 of the rat p53 gene"

gene

/pseudo

15..97

/note="corresponds to exon 2 of the rat p53 gene"

/pseudo

24..1245

/gene="p53"

/pseudo

98..119

/gene="p53"

/note="corresponds to exon 3 of the rat p53 gene"

/pseudo

120..358

/gene="p53"

/note="corresponds to exon 4 of the rat p53 gene"

/pseudo

359..539

/gene="p53"

/note="corresponds to exon 5 of the rat p53 gene"

/pseudo

540..640

/gene="p53"

/note="corresponds to exon 6 of the rat p53 gene"

Mar 25 02:40

US-08-644-289-2.rgc

12

exon /pseudo
641..750
/gene="p53"
/note="corresponds to exon 7 of the rat p53 gene"

exon /pseudo
751..887
/gene="p53"
/note="corresponds to exon 8 of the rat p53 gene"

exon /pseudo
888..961
/gene="p53"
/note="corresponds to exon 9 of the rat p53 gene"

exon /pseudo
962..1153
/gene="p53"
/note="corresponds to exon 10 of the rat p53 gene"

exon /pseudo
1154..>1273
/note="corresponds to a part of exon 11 of the rat p53 gene"

BASE COUNT 328 a 356 c 306 g 283 t
ORIGIN

Query Match 63.2%; Score 24; DB 70; Length 1273;
Best Local Similarity 85.3%; Pred. No. 2.51e-05;

Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 996 aggccttggaatgaagatgccctgcgcaga 1029
||||| || ||||| ||||| ||||| |||||
Oy 5 AGGCTTAGAGTTAAGAGTCCGCTACAGA 38

RESULT 12

LOCUS RNP53 1627 bp RNA ROD 12-SEP-1993

DEFINITION Rat mRNA for nuclear oncoprotein p53.

ACCESSION X13058

NID g56828

KEYWORDS oncoprotein p53.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires; Rodentia;
Sciurognathi; Myomorpha; Muridae; Murinae; Rattus.

1 (bases 1 to 1627)

Sousi, T.

Direct Submision

Submitted (26-SEP-1988) to the EMBL/GenBank/DBJ databases. Sousi
T., Universite Pierre et Marie Curie, Unite d'Oncologie
Moleculaire, IIRC - CNRS, BP 08 94802 Villejuif, France

REFERENCE 2 (bases 1 to 1627)

Sousi, T., Caron de Fromental, C., Breugnot, C. and May, E.
Nucleotide sequence of a cDNA encoding the rat p53 nuclear
oncoprotein

AUTHORS

TITLE

JOURNAL

Nucleic Acids Res. 16 (23), 11384 (1988)

MEDLINE 89083585

COMMENT Data kindly reviewed (09-Feb-1989) by Sousi T.

FEATURES

source Location/Qualifiers

1..1627

/organism="Rattus norvegicus"

/cell_line="Py T21"

CDS 24..1199

```

/note="nuclear protein p53 (AA 1 - 391)"
/db_xref="P1D:956829"
/translation="MEDSOSQMSIELEP1SQETFSCLMKLIPDDILPTTATGSPNSME
DLF1PQVAVLEEGPEALQVSAAPQEPCTAPAPVAPASATPPWPLSSSVSOKTQX
GNYGHFLGELQSTGKAKSVCTQTSISLAKLFLQALATCPVQJMWTSPTPTRVRMAI
YKRSQNMTEVHRCCHERCSGDGLAPQHLIRVGNPYAEYLDDQTRFASVWVY
EPPEDGSDYTTIRHYKNSCMGGMNRPILITITLEDSSGNLIGDSEYRVACAP
GMDRTREENPFRKEHCEPTELPGSARALPTSTSSSPQKKKPLDGEYFTLLKNGE
RPEHREIMDALEIKDAALAESGDSRASHSTPFRKGQSTSHKKPVIKRVGPDSQ"
BASE COUNT      314 a      486 c      380 g      387 t
ORIGIN
Query Match      63.2%, Score 24; DB 69; Length 1627;
Best Local Similarity 85.3%; Pred. No. 2,51e-05;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db      1054 aggccttgaattaaggatgccgtgtgcgca 1087
||||||| || ||||||||| |||| |||
QY      5 AGCGCTTAGACTTAAGATGCCATGCTACACA 38

RESULT 13
ID      RNPSUEDO      standard; DNA; R0D; 1447 BP.
AC      L12046;
DT      16-MAR-1993 (Rel. 35, Created)
DT      14-APR-1996 (Rel. 47, Last updated, Version 3)
DE      Rattus norvegicus p53 (Pc-111) pseudogene, partial ORF.
KW      tumor suppressor.
OS      Rattus norvegicus (rat)
OC      Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC      Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
NC      [1]
RP      1-1447
RA      Weghorst C.M., Buzard G.S., Calvert R.J., Hulla J.E., Rice J.M.;
RT      "Cloning and sequence of a processed p53 pseudogene from rat: a
RT      potential source of false 'mutations' in PCR fragments of tumor
RT      DNA";
RL      Gene 166:317-322(1995).
CC      NCBI gi: 206472
FH      Key      location/Qualifiers
FH      FH      1..1447
FH      source      /organism="Rattus norvegicus"
FH      FT      /strain="Fischer 344"
FH      FT      /sex="male"
FH      FT      /tissue_type="liver"
FH      exon      1..75
FH      FT      /note="this region of the pseudogene corresponds to exon 2
FH      FT      of the rat p53 cDNA, putative"
FH      FT      2..4
FH      FT      /note="this site within the pseudogene corresponds to the
FH      FT      ATG translation start site of the rat p53 cDNA, putative"
FH      FT      76..97
FH      FT      /note="this region of the pseudogene corresponds to exon 3
FH      FT      of the rat p53 cDNA, putative"
FH      exon      98..315
FH      FT      /note="this region of the pseudogene corresponds to exon 4
FH      FT      of the rat p53 cDNA, putative"
FH      FT      <168..>542
FH      FT      /gene="p53 Pc-111"
FH      FT      /note="tumor suppressor; this region of the pseudogene is
potential open reading frame; putative; NCBI gi: 206473"

```

FT		/codon_start=1
FT		/db_xref="PDB:g206473"
FT		/translation="KVGRKRSKCOLLHRLNLELPRLSLKLNLSQILWSSGLPAVSDNQ
V		
FT		CYVHLTSPKLALPLAGEDPEPCAVWGQLHTSNWLHCACGICITKKKSQHMTWRRCSSH
E		
FT		RCSDCDDQTPPPYPPTPSILSG*
FT	exon	316..498
FT		/gene="p53-PC-II1"
FT		/note="this region of the pseudogene corresponds to exon 5
FT		of the rat p53 cDNA; putative"
FT	exon	499..746
FT		/note="this region of the pseudogene corresponds to exon 6
FT		of the rat p53 cDNA; putative"
FT	exon	747..882
FT		/note="this region of the pseudogene corresponds to exon 7
FT		of the rat p53 cDNA; putative"
FT	exon	883..937
FT		/note="this region of the pseudogene corresponds to exon 8
FT		of the rat p53 cDNA; putative"
FT	exon	938..1031
FT		/note="this region of the pseudogene corresponds to exon 9
FT		of the rat p53 cDNA; putative"
FT	exon	1032..1447
0		/note="this region of the pseudogene corresponds to exon 1
FT		of the rat p53 cDNA; putative"
FT	misc signal	1107..1109
FT		/note="this site within the pseudogene corresponds to the
FT		translation stop site of the rat p53 cDNA; putative"
SQ	Sequence	1447 BP; 366 A; 388 C; 336 G; 357 T; 0 other;
<hr/>		
Query Match	52.6%; Score 20; DB 12; Length 1447;	
Best Local Similarity	88.5%; Pred. No. 3.66e-07;	
Matches	23; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
<hr/>		
Db	976 gaattaagatgacatcgtcgaga	1001
Qy		
	13 GAGTTAAGATGCCCGCATCTCAAGA	38
<hr/>		
RESULT	14	
LOCUS	RATPSEUDO	1447 bp DNA ROD 07-JUN-1993
DEFINITION	Rattus norvegicus p53 tumor suppressor pseudogene, partial ORF.	
ACCESSION	L12046	
NID	q206472	
KEYWORDS	tumor suppressor.	
SOURCE	Rattus norvegicus (strain Fisher 344) male liver DNA.	
ORGANISM	Rattus norvegicus	
	Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;	
	Eutheria; Rodentia; Myomorpha; Muridae; Murinae.	
REFERENCE	1 (bases 1 to 1447)	
AUTHORS	Meghorts,C.M., Buzard,G.S., Hulla,J.E., Calvert,R.J. and Rice,J.M.	
TITLE	Nucleotide sequence of a rat p53 pseudogene; a representative of	
JOURNAL	the family of p53 pseudogenes in the rat genome	
FEATURES	Unpublished (1993)	
source	Location/Qualifiers	
	1..1447	
	/organism="Rattus norvegicus"	
	/strain="Fisher 344"	
	/sequenced_mol="DNA"	
	/sex="male"	
	/tissue_type="liver"	
exon	1..75	

Mar 25 02:40

US-08-644-289-2.ngc

15

misc_signal /note="this region of the pseudogene corresponds to exon 2 of the rat p53 cDNA; putative"
2..4 /note="this site within the pseudogene corresponds to the ATG translation start site of the rat p53 cDNA; putative"
76..97 /note="this region of the pseudogene corresponds to exon 3 of the rat p53 cDNA; putative"
98..315 /note="this region of the pseudogene corresponds to exon 4 of the rat p53 cDNA; putative"
CDS <168..>542 /note="this region of the pseudogene is a potential open reading frame; putative"
/codon_start=1 /db_xref="PID:g206473"
/translation="KVQRKRSQQLPHRNLELPISLKNLSQIMSSGLPAYSNDQVCYVHLPSPKLAIPAGEDMPCAWQCQLHTSNMHLCAHGILYKKSQHMTEVMRCSHHERCSDGDDQTPPPYPTPSIISG"
exon 316..498 /note="this region of the pseudogene corresponds to exon 5 of the rat p53 cDNA; putative"
exon 499..746 /note="this region of the pseudogene corresponds to exon 6 of the rat p53 cDNA; putative"
exon 747..882 /note="this region of the pseudogene corresponds to exon 7 of the rat p53 cDNA; putative"
exon 883..937 /note="this region of the pseudogene corresponds to exon 8 of the rat p53 cDNA; putative"
exon 938..1031 /note="this region of the pseudogene corresponds to exon 9 of the rat p53 cDNA; putative"
exon 1032..1447 /note="this region of the pseudogene corresponds to exon 10 of the rat p53 cDNA; putative"
misc_signal 1107..1109 /note="this site within the pseudogene corresponds to the translation stop site of the rat p53 cDNA; putative"
BASE COUNT 366 a 388 c 336 g 357 t
ORIGIN
Query Match 52.6%; Score 20; DB 69; Length 1447;
Best Local Similarity 88.5%; Pred. No. 3.66e-02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Db 976 gaattaagatgacatgctgcaga 1001
Qy 13 GACCTTAAGATGCCATCCTACACA 38

RESULT 15
LOCUS RATTSEUDO 1447 bp DNA ROD 12-APR-1996
DEFINITION Rattus norvegicus p53 (PG-III) pseudogene, partial ORF.
ACCESSION L12046
NID g206472
KEYWORDS tumor suppressor.
SOURCE Rattus norvegicus (strain Fischer 344) DNA.
ORGANISM Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1447)

Mar 25 02:40

US-08-644-289-2.ngc

16

AUTHORS Meghfort, C.M., Buzard, G.S., Calvert, R.J., Hulla, J.E. and Rice, J.M.
TITLE Cloning and sequence of a processed p53 pseudogene from rat: a potential source of false 'mutations' in PCR fragments of tumor DNA
JOURNAL Gene 166 (2), 317-322 (1995)
MEDLINE 96125211
COMMENT NCBI gi: 206472
FEATURES Location/Qualifiers
source 1..1447
/organism="Rattus norvegicus"
/strain="Fischer 344"
/sex="male"
/tissue_type="liver"
1..75
/note="this region of the pseudogene corresponds to exon 2 of the rat p53 cDNA; putative"
2..4
/note="this site within the pseudogene corresponds to the ATG translation start site of the rat p53 cDNA; putative"
76..97
/note="this region of the pseudogene corresponds to exon 3 of the rat p53 cDNA; putative"
98..315
/note="this region of the pseudogene corresponds to exon 4 of the rat p53 cDNA; putative"
CDS <168..>542 /gene="p53 PG-III"
/note="tumor suppressor; this region of the pseudogene is a potential open reading frame; putative; NCBI gi: 206473"
/codon_start=1 /db_xref="PID:g206473"
/translation="KVQRKRSQQLPHRNLELPISLKNLSQIMSSGLPAYSNDQVCYVHLPSPKLAIPAGEDMPCAWQCQLHTSNMHLCAHGILYKKSQHMTEVMRCSHHERCSDGDDQTPPPYPTPSIISG"
exon 316..498 /gene="p53 PG-III"
exon 499..746 /note="this region of the pseudogene corresponds to exon 5 of the rat p53 cDNA; putative"
exon 747..882 /note="this region of the pseudogene corresponds to exon 6 of the rat p53 cDNA; putative"
exon 883..937 /note="this region of the pseudogene corresponds to exon 7 of the rat p53 cDNA; putative"
exon 938..1031 /note="this region of the pseudogene corresponds to exon 8 of the rat p53 cDNA; putative"
exon 1032..1447 /note="this region of the pseudogene corresponds to exon 9 of the rat p53 cDNA; putative"
exon 1107..1109 /note="this region of the pseudogene corresponds to exon 10 of the rat p53 cDNA; putative"

misc_signal 1107..1109 /note="this site within the pseudogene corresponds to the translation stop site of the rat p53 cDNA; putative"
BASE COUNT 366 a 388 c 336 g 357 t
ORIGIN

Query Match 52.6%; Score 20; DB 91; Length 1447;
Best Local Similarity 88.5%; Pred. No. 3.66e-02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Db 976 gaattaagatgacatgctgcaga 1001

Mar 25 02:40

US-08-644-289-2.rge

17

0Y 13 GAGTTAAGGATGCCATGCTACAGA 38
|| ||||| ||||| ||||| |||||

Search completed: Tue Mar 25 02:42:11 1997
Job time : 64 secs.

US-08-644-289-2.rst

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPorch_mn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Mar 25 02:43:09 1997; Maspar time 59.14 Seconds

Tabular output not generated. 283.363 Million cell updates/sec

Title: >US-08-644-289-2

Description: (1-38) from US08644289.seq

Perfect Score: 38

N.A. Sequence: 1 AGTCAGCCCTTAGAGTAAAGATGCCCATCTCTACACA 38

Comp: TCAGTCCGAATCTCAATTCCTACGGGACGATGCT

Scoring table: TABLE default

Gap 10

Mmatch STD : Dbase 0; Query 0

Searched: 630489 segs, 220513910 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

EST-STS

1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8

9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14

15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20

21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26

27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32

33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38

39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44

45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50

51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56

57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62

63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68

69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74

75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80

81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86

87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92

93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98

99:EST99

EST-STS-TWO

100:EST100 101:EST101 102:EST102 103:EST103 104:EST104

105:EST105 106:EST106 107:EST107 108:EST108 109:EST109

110:EST110 111:EST111 112:EST112 113:EST113 114:EST114

115:EST115 116:EST116 117:EST117 118:EST118 119:EST119

120:EST120 121:EST121 122:EST122 123:EST123 124:EST124

125:EST125 126:EST126 127:EST127 128:EST128 129:EST129

Statistics: Mean 7.360; Variance 1.481; scale 4.970

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	18	47.4	368	25	H61357	9.68e-05
2	17	44.7	323	36	H97804	1.89e-03
3	16	42.1	187	39	HSC05H051	3.30e-02
4	16	42.1	277	93	R74999	3.30e-02
5	16	42.1	293	127	T96185	3.30e-02
6	16	42.1	311	52	N20004	3.30e-02
7	16	42.1	325	138	MUSPROBER	3.30e-02
8	16	42.1	337	44	HSC3BH011	3.30e-02
9	16	42.1	364	40	HSC17C011	3.30e-02
10	16	42.1	367	164	HS637121	3.30e-02
11	16	42.1	367	88	R57637	3.30e-02
12	16	42.1	367	152	R57637	3.30e-02
13	16	42.1	367	52	MUS94E02	3.30e-02
14	16	42.1	373	173	HS2902B5	3.30e-02
15	16	42.1	373	135	HS2902B5	3.30e-02
16	16	42.1	394	121	T74343	3.30e-02
17	16	42.1	419	36	H97533	3.30e-02
18	16	42.1	423	74	R09439	3.30e-02
19	16	42.1	424	91	R67886	3.30e-02
20	16	42.1	448	114	T47524	3.30e-02
21	16	42.1	451	163	HS554327	3.30e-02
22	16	42.1	451	145	N98554	3.30e-02
23	16	42.1	467	83	R41287	3.30e-02
24	16	42.1	470	83	R41628	3.30e-02
25	16	42.1	503	96	R84874	3.30e-02
26	16	42.1	515	120	T70346	3.30e-02
27	16	42.1	521	57	N63043	3.30e-02
28	16	42.1	615	54	N25914	3.30e-02
29	15	39.5	158	142	N84100	5.03e-01
30	15	39.5	256	108	T24206	5.03e-01
31	15	39.5	260	173	DM8903T	5.03e-01
32	15	39.5	275	16	H32057	5.03e-01
33	15	39.5	280	72	R03847	5.03e-01
34	15	39.5	287	108	T24206	5.03e-01
35	15	39.5	291	104	T04898	5.03e-01
36	15	39.5	329	166	HS49334	5.03e-01
37	15	39.5	360	3	CEIK040GYF	5.03e-01
38	15	39.5	413	146	M01003	5.03e-01
39	15	39.5	415	127	T93275	5.03e-01
40	15	39.5	433	125	T86880	5.03e-01
41	15	39.5	441	106	T14921	5.03e-01
42	15	39.5	448	131	G07549	5.03e-01
					human STS SHGC-3102 c	5.03e-01

Mar 25 02:42

US-08-644-289-2.rst

3

43 15 39.5 475 68 N71069 za8ec10.s1 Homo sapie 5.03e-01
c 44 15 39.5 487 114 T46248 9511 Arabidopsis thal 5.03e-01
c 45 15 39.5 537 76 R15131 yf88c03.r1 Homo sapie 5.03e-01

ALIGNMENTS

RESULT

1

LOCUS H61357 368 bp mRNA EST 06-OCT-1995
DEFINITION y41406.r1 Homo sapiens cDNA clone 23638 5' similar to
gb:XA54156_rnal CELLULAR TUMOR ANTIGEN P53 (HUMAN);.

ACCESSION

H61357

NID

g1014189

KEYWORDS

EST.
clone=23638 primer=M13RP1 library=Soares ovary tumor NBH07
vector=pf7r3D (Pharmacia) with a modified polylinker host=DH108
(ampicillin resistant) Restrict=Not I Restrict=Eco RI Female. 1st
strand cDNA was primed with a Not I - oligo (dT) primer [5'
TGTACCAATCTGACGTGCGAGCGCGGTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pf7r3 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM

Homo sapiens

REFERENCE

Homo sapiens

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

TITLE

The WashU-Merck EST Project

JOURNAL

Unpublished (1995)

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 284
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

source 1..368
/organism="Homo sapiens"
/clone="23638"

FEATURES

source

BASE COUNT

94 a 99 c 109 g 61 t 5 others

ORIGIN

mRNA

Query Match

47.4%; Score 18; DB 25; Length 368;
Best Local Similarity 79.3%; Pred. No. 9.68e-05;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db

323 aggccttggaactnaagatgctcagct 351

LOCUS

0y

5 AGGCTTGAAGTTAAGAGCATCCATCCT 33

Mar 25 02:42

US-08-644-289-2.rst

4

RESULT 2
LOCUS H97804 323 bp mRNA EST 12-DEC-1995
DEFINITION yw02f03.s1 Homo sapiens cDNA clone 251069 3' similar to contains LI
repetitive element ;.

ACCESSION

H97804

NID

g1118689

KEYWORDS

EST.

human clone=251069 primer=m13 -40 forward library=Soares melanocyte
2bhm vector=pf7r3D (Pharmacia) with a modified polylinker
host=DH108 (ampicillin resistant) Restrict=Not I Restrict=Eco RI Male.
1st strand cDNA was primed with a Not I - oligo (dT) primer
[5'-TGTACCAATCTGACGTGCGAGCGCGGTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pf7r3 vector (Pharmacia). Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.

ORGANISM

Homo sapiens

REFERENCE

Homo sapiens

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

TITLE

The WashU-Merck EST Project

JOURNAL

Unpublished (1995)

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 240
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

source 1..323
/organism="Homo sapiens"
/clone="251069"
/note="human"

FEATURES

source

BASE COUNT

139 a 66 c 55 g 61 t 2 others

ORIGIN

mRNA

Query Match

44.7%; Score 17; DB 36; Length 323;
Best Local Similarity 94.7%; Pred. No. 1.89e-03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db

292 tagaatggcattcctta 310

LOCUS

Cp

34 TAGCATGGCATCCTTTAA 16

RESULT 3
LOCUS HSC05H051 187 bp RNA EST 06-NOV-1994

Mar 25 02:42

US-08-644-289-2.rst

9

Query Match

42.1%; Score 16; DB 138; Length 325;

Best Local Similarity 94.4%; Pred. No. 3.30e-02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 48 aagatgcctctgctaca 65

|||||

Qy 19 AAGATGCCATCGCTACA 36

RESULT

8

LOCUS HSC3RH011 337 bp RNA EST 21-SEP-1995

DEFINITION H. sapiens partial cDNA sequence; clone c-3bh01.

ACCESSION F12586

NID g708585

KEYWORDS partial cDNA sequence; transcribed sequence fragment.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 337)

AUTHORS Genexpress.

TITLE Direct Submission

JOURNAL Submitted (19-JAN-1995) Genethon, B.P. 60, 91002 Evry Cedex France

and Genetique Molculaire et Biologie du developpement, CNRS UPR420

B.P. 8, 94801 Villejuif Cedex France. E-mail: genexpress@genethon.fr

REFERENCE 2 (bases 1 to 337)

AUTHORS Genexpress.

TITLE The Genexpress cDNA program

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 337)

AUTHORS Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C.,

Devignes, M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lamy, B.,

Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,

Sebastiant-Kabakchis, C. and Tessier, A.

IMAGE: molecular integration of the analysis of the human genome

and its expression

C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)

JOURNAL 95277534

COMMENT Cloning method: total mRNA was oligo-(dT) primed and directionally

cloned 5' -> 3' into the HindIII -> NotI sites of the lambdaID BA

vector;

Sequencing method: single read, full automatic;

Primer: MJ3_reverse

cDNA sequence collinear to mRNA

Stretch removed: nothing

Normalization method: Bento Soares, P.N.A.S. 91:9228-9232 (1994);

Genexpress library_id: Cj

Genexpress sequence_id: y1c-3bh01.

location/Qualifiers

FEATURES

source

1..337

/organism="Homo sapiens"

/isolate="muscular atrophy patient"

/clone_lib="normalized infant brain cDNA from B.Soaes,

Psychiatry Dept. Columbia University USA"

/sex="female"

/tissue_type="total brain"

/dev_stage="3 months old"

BASE COUNT 106 a 45 c 62 g 119 t 5 others

ORIGIN

Query Match

42.1%; Score 16; DB 44; Length 337;

Best Local Similarity 83.3%; Pred. No. 3.30e-02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Mar 25 02:42

US-08-644-289-2.rst

10

Db 195 tagaatggatccttgaactata 218

|||||

Cp 34 TAGCATGGCATCCTTTAACTCTA 11

RESULT

9

LOCUS HSC17C011 364 bp RNA EST 06-NOV-1994

DEFINITION H. sapiens partial cDNA sequence; clone c-17c01.

ACCESSION Z43246

NID g570831

KEYWORDS partial cDNA sequence; transcribed sequence fragment.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; mitochondrial eukaryotes; Metazoa; Eumycota group;

Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;

Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;

Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;

Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 364)

AUTHORS Genexpress.

TITLE Direct Submission

JOURNAL Submitted (24-OCT-1994) to the EMBL/Genbank/DBJ databases.

Genethon, B.P. 60, 91002 Evry Cedex France and Genetique

Moleculaire et Biologie du developpement, CNRS UPR420 B.P. 8, 94801

Villejuif Cedex France. E-mail: genexpress@genethon.fr

REFERENCE 2 (bases 1 to 364)

AUTHORS Genexpress.

TITLE The Genexpress cDNA program

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 364)

AUTHORS Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C.,

Devignes, M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lamy, B.,

Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,

Sebastiant-Kabakchis, C. and Tessier, A.

IMAGE: Integrated molecular analysis of the human genome and its

expression

C. R. Acad. Sci., III, Sci. Vie 318, 263-272 (1995)

JOURNAL Clone library from B.Soaes, Psychiatry Dept. Columbia University

COMMENT USA;

Cloning method: total mRNA was oligo-(dT) primed and directionally

cloned 5' -> 3' into the HindIII -> NotI sites of the lambdaID BA

vector;

Sequencing method: single read, full automatic;

Primer: MJ3_reverse

cDNA sequence collinear to mRNA

Stretch removed: nothing

Normalization method: Bento Soares, P.N.A.S. in press;

Genexpress library_id: Cj

Genexpress sequence_id: y1c-17c01;

location/Qualifiers

FEATURES

No significant homology found with :

genbank release 81 swissprot release 28.

location/Qualifiers

1..364

/organism="Homo sapiens"

/dev_stage="3 months old"

/isolate="muscular atrophy patient"

/tissue_type="total brain"

/clone_lib="normalized infant brain cDNA"

BASE COUNT 110 a 51 c 69 g 131 t 3 others

ORIGIN

Mar 25 02:42

US-08-644-289-2.rst

11

Query Match 42.1%; Score 16; DB 40; Length 364;
Best Local Similarity 83.3%; Pred. No. 3.30e-02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 195 taagatggtacctaactata 218
||| ||||| ||||| ||||| ||
Cp 34 TAGCATGGCATCTTTACTCTA 11

RESULT 10
ID HS631121 standard; RNA; EST; 367 BP.
AC R57637;
DT 29-MAY-1995 (Rel. 43, Created)
DT 04-MAY-1996 (Rel. 47, Last updated, Version 3)
DE F4425 Fetal heart Homo sapiens cDNA clone F4425 5' end similar to
DE Myosin alkali light chain.
KM EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-367
RA Huang D.M., Fung Y.W., Wang R.X., Laurensen C.M., Ng S.H.,
RA Lam W.Y., Tsui K.W., Fung K.P., Waye M., Lee C.Y., Liew C.C.;
RT "Analysis of expressed sequence tags from a fetal human heart cDNA
RT library";
RL Genomics 30:293-298(1995).
CC Contact: Liew CC Molecular Cardiology University of Toronto Bantling
CC Institute, 100 College St., Toronto, Ontario, M5G1L5 Tel:
CC 4169788758 Fax: 4169785650 Email: liewc@utcc.utoronto.ca Seq
CC primer: GGTCGCAGCACTCTGAGCC. NCBI gi: 827695
FH Key Location/Qualifiers
FH
FT source 1..367
FT /organism="Homo sapiens"
FT /note="Vector: lambda gt22; Site 1: NotI; Site 2: SalI;
FT mRNA was purified from human fetal hearts (10-12 weeks).
FT cDNA was constructed using a NotI-Oligo dT adaptor-primer.
FT SalI adaptors were ligated, followed by digestion with
FT NotI, for directioncloning into predigested lambda gt22.
FT Method is described in J. Mol. Cell. Cardiol. (1994) 26,
FT 1329-1333)*
FT /clone="F4425"
FT /clone.lib="fetal heart"
FT /lab host="E. coli Y1090"
FT mRNA
FT <1..>367
SQ Sequence 367 BP; 98 A; 75 C; 117 G; 77 T; 0 other;

Query Match 42.1%; Score 16; DB 164; Length 367;
Best Local Similarity 83.3%; Pred. No. 3.30e-02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 18 cagaccgtagatttaagagcc 41
||| || ||||| ||||| ||||
Qy 4 CAGGCTTAGACTTAAGCATGCC 27

RESULT 11
LOCUS R57637 367 bp mRNA EST 23-MAY-1995
DEFINITION F4425 Homo sapiens cDNA clone F4425 5' end similar to Myosin alkali
light chain.
ACCESSION R57637
NTD g827695

Mar 25 02:42

US-08-644-289-2.rst

12

KEYWORDS EST.
SOURCE human clone=F4425 library=Fetal heart vector=Lambda gt22 host=E.
coli Y1090 primer=GGTCGCAGCACTCTGAGCC Raltel=NotI Raltel2=SalI
mRNA was purified from human fetal hearts (10-12 weeks). cDNA was
constructed using a NotI-Oligo dT adaptor-primer. SalI adaptors
were ligated, followed by digestion with NotI, for direction
cloning into predigested lambda gt22. Method is described in J.
Mol. Cell. Cardiol. (1994) 26, 1329-1333).

ORGANISM Homo sapiens
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 367)
Huang D.M., Fung Y.W., Wang R.X., Laurensen C.M., Ng S.H.,
Lam W.Y., Tsui K.W., Fung K.P., Waye M., Lee C.Y. and Liew C.C.
Analysis of Expressed Sequence Tags (ESTs) from Fetal and Adult
Heart cDNA Libraries
Unpublished (1995)

JOURNAL
COMMENT

CONTACT: Liew CC
Molecular Cardiology
University of Toronto
Bantling Institute, 100 College St., Toronto, Ontario, M5G1L5
TEL: 4169788758
Fax: 4169785650
Email: liewc@utcc.utoronto.ca.
Location/Qualifiers
1..367
/organism="Homo sapiens"
/clone="F4425"
/note="human"

BASE COUNT 98 a 75 c 117 g 77 t
ORIGIN

Query Match 42.1%; Score 16; DB 88; Length 367;
Best Local Similarity 83.3%; Pred. No. 3.30e-02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 18 cagaccgtagatttaagagcc 41
||| || ||||| ||||| ||||
Qy 4 CAGGCTTAGACTTAAGCATGCC 27

RESULT 12
LOCUS R57637 367 bp mRNA EST 02-MAY-1996
DEFINITION F4425 Fetal heart Homo sapiens cDNA clone F4425 5' end similar to
Myosin alkali light chain.
ACCESSION R57637
NTD g827695
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 367)
Huang D.M., Fung Y.W., Wang R.X., Laurensen C.M., Ng S.H.,
Lam W.Y., Tsui K.W., Fung K.P., Waye M., Lee C.Y. and Liew C.C.
Analysis of expressed sequence tags from a fetal human heart cDNA
library
Genomics 30 (2), 293-298 (1995)

JOURNAL
COMMENT
Contact: Liew CC
Molecular Cardiology

Mar 25 02:42

US-08-644-289-2.rst

15

RESULT 15

LOCUS HS290285 373 bp DNA STS 22-MAR-1996

DEFINITION H.sapiens (D2S2334) DNA segment containing (CA) repeat; clone AFW2902b5; single read.

ACCESSION 251281

NID q1232581

KEYWORDS CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA; microsatellite marker; repeat polymorphism; STS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 373)

TITLE Weissenbach, J.

JOURNAL Direct Submission

REFERENCE Submitted (01-SEP-1995) Genethon, B.P. 60, 91002 Evry Cedex France.

AUTHORS E-mail: Jean.Weissenbach@genethon.fr

TITLE 2 (bases 1 to 373)

JOURNAL Dlb, C., Faure, S., Fitzames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J. and Weissenbach, J.

COMMENT A comprehensive genetic map of the human genome based on 5,264 microsatellites

FEATURES Nature 380, 152-154 (1996)

source full automatic.

location/Qualifiers

1..373

/organism="Homo sapiens"

/note="cloning vector is M13mp18"

/cell_line="CEPH 134702"

/clone_lib="genomic DNA"

/chromosome="2"

BASE COUNT 132 a 87 c 53 g 95 t 6 others

ORIGIN

Query Match 42.1%; Score 16; DB 135; length 373;

Best Local Similarity 78.6%; Pred. No. 3.30e-02;

Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 80 catggccaccttaataaagcctg 107

||||||| ||||||| ||| |||||||

Cp 31 CATGGCCTCTTAACTTAAGCCCTG 4

Search completed: Tue Mar 25 02:44:17 1997
Job time : 68 secs.

INVENTED

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_mn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Mar 25 02:47:33 1997; MaapPar time 8.73 Seconds
279.817 Million cell updates/sec

Tabular output not generated.

Title: >US-08-644-289-3
Description: (1-28) from US08644289.seq
Perfect Score: 28
N.A. Sequence: 1 ACTCGAATTCATTGGGACCATCTGGCT 28
Comp: TCAGCTTAAGTAACCTCGTGAAGACCGA

Scoring table: TABLE default
Gap 10

Nmatch STD : Dbase 0; Query 0

Searched: 113505 seqs, 43611913 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

n-geneseq25
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22

Statistics: Mean 5.755; Variance 3.299; scale 1.744

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB	ID	Description	Pred. No.
1	28	100.0	28	21	T10487 Human wild-type p53 g	7.90e-07
2	23	82.1	91	9	Q51746 Oligonucleotide probe	8.59e-04
3	20	71.4	91	9	Q51746 Oligonucleotide probe	4.74e-02
4	16	57.1	184	6	Q39756 Expressed Sequence Ta	7.34e+00
5	16	57.1	184	8	Q59168 Human brain Expressed	7.34e+00
6	16	57.1	280	8	Q61255 Human brain Expressed	7.34e+00
7	16	57.1	303	6	Q39930 Expressed Sequence Ta	7.34e+00
8	16	57.1	303	8	Q59342 Human brain Expressed	7.34e+00

c	9	16	57.1	342	8	Q61318 Human brain Expressed	7.34e+00
	10	16	57.1	406	6	Q39880 Expressed Sequence Ta	7.34e+00
	11	16	57.1	406	8	Q59292 Human brain Expressed	7.34e+00
c	12	16	57.1	2020	9	Q53209 Human cyclin D3 pseud	7.34e+00
c	13	16	57.1	2022	5	Q31877 Cyclin D3 pseudogene	7.34e+00
c	14	16	57.1	3094	3	N92579 Sequence of the 1.7kb	7.34e+00
c	15	16	57.1	4721	6	Q36395 Benzene dioxygenase g	7.34e+00
c	16	16	57.1	13585	17	T11549 Tumour rejection anti	7.34e+00
c	17	15	53.6	162	9	Q55095 Degenerate oligo for	2.41e+01
c	18	15	53.6	267	3	N50521 Sequence encoding sal	2.41e+01
c	19	15	53.6	801	2	N71120 Sequence encoding pig	2.41e+01
c	20	15	53.6	955	3	N50531 Plasmid pPE603.	2.41e+01
c	21	15	53.6	1767	1	Q06900 Neo-pullulanase gene.	2.41e+01
c	22	15	53.6	2392	15	Q92775 Human thymopoietin-ga	2.41e+01
c	23	15	53.6	2392	18	T13451 Human thymopoietin-ga	2.41e+01
c	24	15	53.6	2392	20	T26961 Thymopoietin-gamma cD	2.41e+01
c	25	15	53.6	2419	19	T16883 Human growth differen	2.41e+01
c	26	15	53.6	3728	10	Q58708 Mouse OSF-5 coding se	2.41e+01
c	27	15	53.6	19012	6	Q36952 HSA gene.	2.41e+01
c	28	14	50.0	204	1	N81164 Base substituted E.co	7.64e+01
c	29	14	50.0	260	8	Q59313 Human brain Expressed	7.64e+01
c	30	14	50.0	377	8	Q59197 Human brain Expressed	7.64e+01
c	31	14	50.0	1140	2	Q05617 Exon 1 of porcine lau	7.64e+01
c	32	14	50.0	1482	2	Q11820 Plasmidogen activator	7.64e+01
c	33	14	50.0	1494	1	Q03369 Sequence encoding hum	7.64e+01
c	34	14	50.0	1990	2	N92451 cDNA insert of vector	7.64e+01
c	35	14	50.0	2581	2	N70974 Sequence of human act	7.64e+01
c	36	14	50.0	2771	4	Q24257 Branching enzyme DNA.	7.64e+01
c	37	14	50.0	3096	2	N71063 Genomic sequence enco	7.64e+01
c	38	14	50.0	3207	3	Q14263 p-meta-1 metastaasis-s	7.64e+01
c	39	14	50.0	3378	12	Q71479 Consensus sequence of	7.64e+01
c	40	14	50.0	3480	5	Q28568 Human thrombin recept	7.64e+01
c	41	14	50.0	4553	4	Q25591 Lipoxigenase structur	7.64e+01
c	42	14	50.0	4922	21	T18877 Human TGL-1 gene.	7.64e+01
c	43	14	50.0	8342	13	Q75209 ALL-1 (acute lymphocy	7.64e+01
c	44	14	50.0	11357	9	Q51024 Human FCER1 beta chai	7.64e+01
c	45	14	50.0	17327	7	Q44278 Serglycin - proteogly	7.64e+01

ALIGNMENTS

RESULT 1
ID T10487 standard; DNA: 28 BP.
AC T10487;
DT 03-OCT-1996 (first entry)
DE Human wild-type p53 gene N-terminal RT-PCR sense primer.
KM Tumour suppressor; p53; alternative splicing; antibody; cancer;
KM proliferative disorder; prognosis; diagnosis; cell cycle;
KM reverse transcriptase polymerase chain reaction; ss.
OS Homo sapiens.
PN CA2150994-A.
PD 15-DEC-1995.
PF 05-JUN-1995; 150994.
PR 14-JUN-1994; US-259612.
PA (HEAL-) HEALTH RES INC.
PI Kulesz-Martin MF;
DR WPI; 96-140137/15.
PT New antibodies specific for alternatively spliced mammalian p53 -
PT useful in diagnosis or prognosis of cancer, and for establishing
PT individual treatment regimes.
PS Disclosure; Page 7; 40pp; English.
CC T10486-T10488 are primers used in a RT-PCR reaction to obtain an
CC N-terminal (nucleotides -111 to 1090) portion of the wild-type p53
CC gene. An epitope from intron 10 of the human p53 gene is used to

RESULT 5
ID 059168 standard; cDNA, 184 BP.
AC 059168;
DT 16-MAR-1994 (first entry)
DE Human brain Expressed Sequence Tag EST00322.
KM Gene transcription product; genetic markers; tagging; in vivo;
transcription; mapping; locations; chromosomes; chromosomal; ss.
OS Homo sapiens.
PM W09316178-A.
PD 19-AUG-1993.
PF 12-FEB-1993; 001294.
PR 12-FEB-1992; US-837195.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Adams MD, Moreno RF, Venter CJ.
DR WPI; 93-272882/34.
PT Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
of most human genes
PS Example 1; Page 132; 500pp; English.
CC The Expressed Sequence Tag was isolated from a human brain cDNA
CC library as part of a large set of ESTs which can be used as markers
CC for human genes transcribed in vivo. They can be used to facilitate
CC tagging of most human genes, for mapping locations of expressed genes
CC on chromosomes, for individual or forensic identification, for mapping
CC locations of disease-associated genes, for identification of tissue
CC type, and for prep. of antisense sequences, probes and constructs.
CC EST00322 has a "poor" coding probability as evaluated using the
CC coding-region prediction program CRM.
CC See also Q59041-Q61440.
SQ Sequence 184 BP; 40 A; 52 C; 45 G; 47 T;

Query Match 57.1%; Score 16; DB 8; Length 184;
Best Local Similarity 94.4%; Pred. No. 7.34e+00;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 163 agccagatggtctcaat 180
|||||
CP 28 ACCCAGCATGTCCTCAAT 11

RESULT 6
ID 061255 standard; DNA; 280 BP.
AC 061255;
DT 16-MAR-1994 (first entry)
DE Human brain Expressed Sequence Tag EST01742.
KM Gene transcription product; genetic markers; tagging; in vivo;
transcription; mapping; locations; chromosomes; chromosomal; ss.
OS Homo sapiens.
PM W09316178-A.
PD 19-AUG-1993.
PF 12-FEB-1993; 001294.
PR 12-FEB-1992; US-837195.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Adams MD, Moreno RF, Venter CJ.
DR WPI; 93-272882/34.
PT Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
of most human genes
PS Example 4; Page 462; 500pp; English.
CC The Expressed Sequence Tag was isolated from a human brain cDNA
CC library as part of a large set of ESTs which can be used as markers
CC for human genes transcribed in vivo. They can be used to facilitate
CC tagging of most human genes, for mapping locations of expressed genes

CC on chromosome, for individual or forensic identification, for mapping
CC locations of disease-associated genes, for identification of tissue
CC type, and for prep. of antisense sequences, probes and constructs.
CC EST01742 has a "poor" coding probability as evaluated using the
CC coding-region prediction program CRM. See also Q59041-Q61440.
SQ Sequence 280 BP; 61 A; 62 C; 49 G; 107 T;

Query Match 57.1%; Score 16; DB 8; Length 280;
Best Local Similarity 94.4%; Pred. No. 7.34e+00;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 44 agccagatggtctcaat 61
|||||
CP 28 ACCCAGCATGTCCTCAAT 11

RESULT 7
ID 039930 standard; DNA; 303 BP.
AC 039930;
DT 20-MAY-1993 (first entry)
DE Expressed Sequence Tag human gene marker EST00312.
KM expressed sequence tag; human genome project; chromosome;
KM human gene sequencing; PCR mapping; somatic cell hybrids;
KM sublocalisation; gene tagging; tissue typing.
OS Synthetic.
PM W09300353-A.
PD 07-JAN-1993.
PF 19-JUN-1992; 005222.
PR 20-JUN-1991; US-716831.
PR 12-FEB-1992; US-837195.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Adams MD, Venter CJ.
DR WPI; 93-036325/04.
PT Particular expressed sequence tags from human cDNA - corresponds
PT to transcription prod. of genes, useful for tagging genes,
PT mapping chromosomes and tissue typing
PS Claim 3; Page 189; 199pp; English.
CC This sequence represents an EST (expressed sequence tag) ESTs are markers
CC for human genes actually transcribed in vivo. Unlike the random genomic
CC DNA sequence tagged sites (STSs), ESTs point directly to expressed genes.
CC The use of ESTs could facilitate the tagging of most expressed human
CC genes within a few years at a fraction of the cost of complete genomic
CC sequencing. Using PCR primers Q39419-Q39580 (sequences designed
CC from the ESTs) sublocalisation of an EST can be achieved with panels of
CC fragments from specific chromosomes or pools of large genomic clones in
CC an analogous manner. This sequence represents EST00312.
SQ Sequence 303 BP; 102 A; 53 C; 67 G; 79 T;

Query Match 57.1%; Score 16; DB 6; Length 303;
Best Local Similarity 94.4%; Pred. No. 7.34e+00;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 259 attagaccatctcgct 276
|||||
CY 11 ATTGGACCATCTCTCGCT 28

RESULT 8
ID 059342 standard; cDNA; 303 BP.
AC 059342;
DT 16-MAR-1994 (first entry)
DE Human brain Expressed Sequence Tag EST00312.
KM Gene transcription product; genetic markers; tagging; in vivo;
transcription; mapping; locations; chromosomes; chromosomal; ss.

OS Homo sapiens.
 PN W09316178-A.
 PD 19-AUG-1993.
 PF 12-FEB-1993; U01294.
 PR 12-FEB-1992; US-837195.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Adams MD, Moreno RF, Venter CJ.
 DR WPI; 93-272882/34.
 PT Enriched oligonucleotides and corresp. sequences - used as
 markers for human genes transcribed in-vivo, facilitate tagging
 of most human genes
 PS Example 1; Page 159; 500bp; English.
 CC The Expressed Sequence Tag was isolated from a human brain cDNA
 library as part of a large set of ESTs which can be used as markers
 for human genes transcribed in vivo. They can be used to facilitate
 tagging of most human genes, for mapping locations of expressed genes
 on chromosomes, for individual or forensic identification, for mapping
 locations of disease-associated genes, for identification of tissue
 CC type, and for prep. of antisense sequences, probes and constructs.
 CC EST00312 has a "poor" coding probability as evaluated using the
 CC coding-region prediction program CDM.
 CC See also Q59041-061440.
 SQ Sequence 303 BP; 102 A; 53 C; 67 G; 79 T;

Query Match 57.1%; Score 16; DB 8; Length 303;
 Best Local Similarity 94.4%; Pred. No. 7.34e+00;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 259 attgagaccatcctgct 276
 ||||| ||||| ||||| |||||
 Oy 11 ATTGGACATCTGGCT 28

RESULT 9
 ID 061318 standard; DNA; 342 BP.
 AC 061318;
 DT 16-MAR-1994 (first entry)
 DE Human brain Expressed Sequence Tag EST01770.
 KM Gene transcription product; genetic markers; tagging; in vivo;
 KM transcription; mapping; locations; chromosomes; chromosomal; ss.
 OS Homo sapiens.
 PN W09316178-A.
 PD 19-AUG-1993.
 PF 12-FEB-1993; U01770.
 PR 12-FEB-1992; US-837195.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Adams MD, Moreno RF, Venter CJ.
 DR WPI; 93-272882/34.
 PT Enriched oligonucleotides and corresp. sequences - used as
 markers for human genes transcribed in-vivo, facilitate tagging
 of most human genes
 PS Example 4; Page 472; 500bp; English.
 CC The Expressed Sequence Tag was isolated from a human brain cDNA
 library as part of a large set of ESTs which can be used as markers
 for human genes transcribed in vivo. They can be used to facilitate
 tagging of most human genes, for mapping locations of expressed genes
 on chromosomes, for individual or forensic identification, for mapping
 locations of disease-associated genes, for identification of tissue
 CC type, and for prep. of antisense sequences, probes and constructs.
 CC EST01770 has a "poor" coding probability as evaluated using the
 CC coding-region prediction program CDM. See also Q59041-061440.
 SQ Sequence 342 BP; 60 A; 100 C; 88 G; 93 T;

Query Match 57.1%; Score 16; DB 8; Length 342;

Best Local Similarity 94.4%; Pred. No. 7.34e+00;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 163 acccagatgctccaat 180
 ||||| ||||| ||||| |||||
 Cp 28 AGCCAGATGCTCCCAAT 11

RESULT 10
 ID 039880 standard; DNA; 406 BP.
 AC 039880;
 DT 20-MAY-1993 (first entry)
 DE Expressed Sequence Tag human gene marker EST00337.
 KM expressed sequence tag; human genome project; chromosome;
 KM human gene sequencing; PCR mapping; somatic cell hybrids;
 KM sublocalisation; gene tagging; tissue typing.
 OS Synthetic.
 PN W09300353-A.
 PD 07-JAN-1993.
 PF 19-JUN-1992; U05222.
 PR 20-JUN-1991; US-716831.
 PR 12-FEB-1992; US-837195.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Adams MD, Venter JC;
 DR WPI; 93-036325/04.
 PT Particular expressed sequence tags from human cDNA - corresponds
 PT to transcription prods. of genes, useful for tagging genes,
 PT mapping chromosomes and tissue typing
 PS Claim 3; Page 171; 199p; English.
 CC This sequence represents an EST (expressed sequence tag) ESTs are markers
 for human genes actually transcribed in vivo. Unlike the random genomic
 CC DNA sequence tagged sites (STSs), ESTs point directly to expressed genes.
 CC The use of ESTs could facilitate the tagging of most expressed human
 CC genes within a few years at a fraction of the cost of complete genomic
 CC sequencing. Using PCR primers Q39419-Q39580 (sequences designed
 CC from the ESTs) sublocalisation of an EST can be achieved with panels of
 CC fragments from specific chromosomes or pools of large genomic clones in
 CC an analogous manner. This sequence represents EST00337.
 SQ Sequence 406 BP; 103 A; 109 C; 114 G; 77 T;

Query Match 57.1%; Score 16; DB 6; Length 406;
 Best Local Similarity 94.4%; Pred. No. 7.34e+00;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 171 attgagaccatcctgct 188
 ||||| ||||| ||||| |||||
 Oy 11 ATTGGACATCTGGCT 28

RESULT 11
 ID Q59292 standard; cDNA; 406 BP.
 AC Q59292;
 DT 16-MAR-1994 (first entry)
 DE Human brain Expressed Sequence Tag EST00337.
 KM Gene transcription product; genetic markers; tagging; in vivo;
 KM transcription; mapping; locations; chromosomes; chromosomal; ss.
 OS Homo sapiens.
 PN W09316178-A.
 PD 19-AUG-1993.
 PF 12-FEB-1993; U01294.
 PR 12-FEB-1992; US-837195.
 PR 12-FEB-1992; US-837195.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Adams MD, Moreno RF, Venter CJ.
 DR WPI; 93-272882/34.

PT Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
PT of most human genes
PS Example 1; Page 151; 500bp; English.
CC The Expressed Sequence Tag was isolated from a human brain cDNA
CC library as part of a large set of ESTs which can be used to facilitate
CC for human gene transcribed in vivo. They can be used to facilitate
CC tagging of most human genes, for mapping locations of expressed genes
CC on chromosomes, for individual or forensic identification, for mapping
CC locations of disease-associated genes, for identification of tissue
CC type, and for prepn. of antisense sequences, probes and constructs.
CC EST00337 has a "poor" coding probability as evaluated using the
CC coding-region prediction program CRM.
CC See also Q59041-Q61440.
SQ Sequence 406 BP; 103 A; 109 C; 114 G; 77 T;

Query Match 57.1%; Score 16; DB 8; Length 406;
Best Local Similarity 94.4%; Pred. No. 7.34e+00;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 171 attgagacatctcgct 188
||||| ||||||| |||||
Qy 11 ATTCGACATCTCGCT 28

RESULT 12
ID 053209 standard; DNA; 2020 BP.
AC Q53209;
DT 22-JUN-1994 (first entry)
DE Human cyclin D3 pseudogene.
KM D-type; mammalian; Cln protein; protein deficiency; cell cycle start;
KM yeast; complement; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1136..1793
FT /tag= a
PN M09324514-A.
PD 09-DEC-1993.
PF 25-MAY-1993; U05000.
PR 26-MAY-1992; US-888178.
PA (MITO-) MITOTIX.
PI Beach DH.
DR WPI; 93-405720/50.
DR P-PSDB; R44806.
PT New D-type mammalian cyclin - replaces Cln-type protein needed
PT for cell start in budding yeast and is detected by antibodies or
PT hybridisation in biological samples to determine abnormal cell
PT division
PS Disclosure; Fig 10; 108bp; English.
CC The sequence is that of a human cyclin D3 pseudogene.
CC Sequence 2020 BP; 553 A; 485 C; 463 G; 519 T;

Query Match 57.1%; Score 16; DB 9; Length 2020;
Best Local Similarity 94.4%; Pred. No. 7.34e+00;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 826 acccagatggtctcaat 843
||||| ||||||| |||||
Cp 28 ACCCAGATGTCCTCAAT 11

RESULT 13
ID Q31877 standard; DNA; 2022 BP.
AC Q31877;

DT 22-APR-1993 (first entry)
DE Cyclin D3 pseudogene.
KM Cyclin D2; D3; lambdaD2-G1; lambdaD3-G5; destruction box; A; B;
KM ubiquitin-dependant; degradation; similarity; D1; E; ss.
OS Homo sapiens.

FH Key Location/Qualifiers
FT CDS 1134..1793

FT /tag= a
FT mutation 1265..1267

FT /tag= b
FT /note= "Nonsense mutation"
FT mutation 1210

FT /tag= c
FT /note= "Frameshift mutation"
FT mutation 1679

FT /tag= d
FT /note= "Frameshift mutation"
FT repeat_unit 1361..1375

FT /tag= e
FT /rpt_type= Direct duplication
FT repeat_unit 1379..1393

FT /tag= f
FT /rpt_type= Direct duplication
FT repeat_unit 1379..1393

FT /tag= g
FT /note= "A to G mutation"
PN M09220796-A.

PD 26-NOV-1992.
PF 18-MAY-1992; U04146.
PR 16-MAY-1991; US-701514.

PA (COLD-) COLD SPRING HARBOR LAB.
PI Beach DH;
DR WPI; 92-415774/50.

DR P-PSDB; R29314.
PT Recombinant mammalian D-type cyclin - replaces a Cln-type protein
PT essential for cell start in budding yeast, its antibodies and
PT probes being useful in detecting D-type cyclin in biological
PT samples

PS Disclosure; Fig 10; 75pp; English.
CC The sequences given in Q31876-77 represents the cyclin D2 and D3
CC pseudogenes. These sequences were identified during the chromosomal
CC localisation of the D2 and D3 genes. These genes were contained
CC within clones lambdaD2-G1 and lambdaD3-G5 respectively. The three
CC novel D-type cyclins disclosed by this invention were shown to encode
CC a similar small size protein ranging from 289-295 amino acids. They
CC lack the "destruction box" identified in the N-terminus of both A-
CC and B-type cyclins which targets them for ubiquitin-dependant
CC degradation. This suggests that the D-type cyclins have evolved a
CC different mechanism to govern their periodic degradation during cell
CC cycle or that they do not undergo destruction. The human cyclin D
CC genes share very high similarity over their entire coding region: 60%
CC between D1 and D2, 60% between D2 and D3 and 52% between D1 and D3.
CC The cyclin D genes are more closely related than the cyclin B genes
CC but are less closely related than the cyclin A and cyclin E genes.
CC See also Q31873-75.

CC Sequence 2022 BP; 554 A; 487 C; 463 G; 518 T;

Query Match 57.1%; Score 16; DB 5; Length 2022;
Best Local Similarity 94.4%; Pred. No. 7.34e+00;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 827 acccagatggtctcaat 844
||||| ||||||| |||||
Cp 28 ACCCAGATGTCCTCAAT 11

```
RESULT 14
ID N92579 standard; DNA; 3094 BP.
AC N92579;
DT 15-MAR-1992 (first entry)
DE Sequence of the 1.7kb cDNA molecule encoding antibodies 7D1, 7D4
DE and 20C6.
KM Vaccine; coccidiosis; poultry; bivalent vaccine; ss.
OS Eimeria.
FH Key
FT CDS Location/Qualifiers
FT /tag= a
FT EP-344808-A.
FT 06-DEC-1989.
FT 02-JUN-1989; 110056.
PR 03-JUN-1988; US-202721.
PA (HOPE ) HOFFMANN-LA ROCHE AG.
PI Allenburger W, Binger MH, Chizzonite RA, Kramer RA,
PI Lomedico PT, McAndrew SJ;
DR WP1; 89-358220/49.
DR P-PsDB; P93706.
PT DNA sequences encoding Eimeria surface antigens - used in
PT pox-virus vectors as a vaccine to protect chicks against
PT coccidiosis
PS Claim 10; Fig 20A-D; 78pp; English.
CC The inventors claim a new protein which comprises one or more
CC immunoreactive and/or antigenic determinants of an Eimeria surface
CC antigen of mol. wt. 28,37,120 or more than 200 kD which specifically
CC binds to one or more monoclonal antibody (Mab) from ATCC HB 9707-9712
CC (see P93703-6). Also new are DNA encoding the protein (see N92576-9),
CC and a vaccine comprising one or more proteins. Vaccine utility can
CC be enhanced by inserting additional genes into the carrier virus
CC (see P91652).
SQ Sequence 3094 BP; 617 A; 834 C; 846 G; 796 T;

Query Match 57.1%; Score 16; DB 3; length 3094;
Best Local Similarity 80.8%; Pred. No. 7.34e+00;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
Db 2917 tgcactcatlgaacctcttgct 2942
||| | ||||| ||| || |||||
Oy 3 TCGAATTCATTCGCGCATCTCGCT 28

RESULT 15
ID Q36385 standard; DNA; 4721 BP.
AC Q36395;
DT 07-JUN-1993 (first entry)
DE Benzene dioxygenase gene region.
KM Benzene; dioxygenase; cybenzene glycol; transformation; BDO; ss.
OS Pseudomonas aeruginosa.
FH Key
FT CDS Location/Qualifiers
FT /tag= a
FT /note= "Benzene dioxygenase gene I"
FT CDS 1912..2553
FT /tag= b
FT /note= "Benzene dioxygenase gene II"
FT CDS 2591..2998
FT /tag= c
FT /note= "Benzene dioxygenase gene III"
FT CDS 3084..3413
FT /tag= d
```

```
FT /note= "Benzene dioxygenase gene IV"
FT CDS 3410..4633
FT /tag= e
FT /note= "Benzene dioxygenase gene V"
FT -10 signal 479..484
FT /tag= f
FT -35 signal 455..460
FT /tag= g
PN J05003785-A.
PD 14-JAN-1993.
PE 30-OCT-1991; 284795.
PR 16-APR-1991; JP-084170.
PA (SAOC ) MERTCAN CORP.
DR WP1; 93-054795/07.
DR P-PsDB; R32085-89.
PT Microbe DNA - contg. benzene di:oxygenase gene, cis:benzene
PT glycol dehydrogenase gene and catechol 2,3-oxygenase gene
PS Claim 11; Page 12-18; 20pp; Japanese.
CC This sequence encodes five regions of a benzene dioxygenase (BDO)
CC polypeptide. Cysbenzene glycol can be produced effectively from
CC benzene using transformants prepared using this fragment.
SQ Sequence 4721 BP; 917 A; 1382 C; 1494 G; 928 T;

Query Match 57.1%; Score 16; DB 6; length 4721;
Best Local Similarity 94.4%; Pred. No. 7.34e+00;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 1524 tgytccaatgaatcga 1541
||||| ||||| ||||| |||||
Cp 20 TCGTCCCATGAATTCGA 3
```

Search completed: Tue Mar 25 02:47:48 1997
Job time : 15 secs.

US-08-644-289-3 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_mn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Mar 25 02:46:12 1997; MasPar time 54.01 Seconds

Tabular output not generated. 426,945 Million cell updates/sec

Title: >US-08-644-289-3

Description: (1-28) from US08644289.seq

Perfect Score: 28

N.A. Sequence: 1 AGTCGATTCTTGGAGCACCCTCGCT 28

Comp: TCAGCTTAAGTACCCCTGCTGAGACCGA

Scoring table: TABLE default

Gap 10

Match STD : Dbase 0; Query 0

Searched: 279077 segs, 411808665 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: embl-new5

1:BCT 2:ETON 3:INV1 4:INV2 5:ORG 6:MAM 7:VRT 8:PLN 9:PRI

10:PRO1 11:PRO2 12:ROD 13:SYN 14:UNC 15:VIR

Database: genbank94

16:BCT1 17:BCT2 18:BCT3 19:BCT4 20:BCT5 21:BCT6 22:BCT7

23:BCT8 24:BCT9 25:INV1 26:INV2 27:INV3 28:INV4 29:INV5

30:INV6 31:INV7 32:MAM1 33:MAM2 34:MAM3 35:VRT1 36:VRT2

37:VRT3 38:VRT4 39:VRT5 40:VRT6 41:VRT7 42:VRT8 43:VRT9

44:VRT10 45:VRT11 46:VRT12 47:VRT13 48:VRT14 49:VRT15 50:VRT16

51:VRT17 52:VRT18 53:VRT19 54:VRT20 55:VRT21 56:VRT22 57:VRT23

58:VRT24 59:VRT25 60:VRT26 61:VRT27 62:VRT28 63:VRT29 64:VRT30

65:VRT31 66:VRT32 67:VRT33 68:VRT34 69:VRT35 70:VRT36 71:VRT37

72:VRT38 73:VRT39 74:VRT40 75:VRT41 76:VRT42 77:VRT43 78:VRT44

79:VRT45 80:VRT46 81:VRT47 82:VRT48 83:VRT49 84:VRT50 85:VRT51

86:VRT52 87:VRT53 88:VRT54 89:VRT55 90:VRT56 91:VRT57 92:VRT58

93:VRT59 94:VRT60 95:VRT61 96:VRT62 97:VRT63 98:VRT64 99:VRT65

100:VRT66 101:VRT67 102:VRT68 103:VRT69 104:VRT70 105:VRT71

106:VRT72 107:VRT73 108:VRT74 109:VRT75 110:VRT76 111:VRT77

112:VRT78 113:VRT79 114:VRT80 115:VRT81 116:VRT82 117:VRT83

118:VRT84 119:VRT85 120:VRT86 121:VRT87 122:VRT88 123:VRT89

124:VRT90 125:VRT91 126:VRT92 127:VRT93 128:VRT94 129:VRT95

130:VRT96 131:VRT97 132:VRT98 133:VRT99 134:VRT100 135:VRT101

136:VRT102 137:VRT103 138:VRT104 139:VRT105 140:VRT106 141:VRT107

142:VRT108 143:VRT109 144:VRT110 145:VRT111 146:VRT112 147:VRT113

148:VRT114 149:VRT115 150:VRT116 151:VRT117 152:VRT118 153:VRT119

154:VRT120 155:VRT121 156:VRT122 157:VRT123 158:VRT124 159:VRT125

160:VRT126 161:VRT127 162:VRT128 163:VRT129 164:VRT130 165:VRT131

166:VRT132 167:VRT133 168:VRT134 169:VRT135 170:VRT136 171:VRT137

172:VRT138 173:VRT139 174:VRT140 175:VRT141 176:VRT142 177:VRT143

178:VRT144 179:VRT145 180:VRT146 181:VRT147 182:VRT148 183:VRT149

184:VRT150 185:VRT151 186:VRT152 187:VRT153 188:VRT154 189:VRT155

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	21	75.0	536	65	MUSANTP53	Mouse tumor antigen p	4.30e-03
2	21	75.0	536	64	MUSANTP53	Mouse tumor antigen p	4.30e-03
3	21	75.0	554	64	MUSANTP53	Mouse tumor antigen p	4.30e-03
4	21	75.0	1285	67	MUSANTP53	Mouse p53 mRNA, compl	4.30e-03
5	21	75.0	1772	67	MUSANTP53	Mouse p53 cellular tu	4.30e-03
6	21	75.0	1773	64	MUSANTP53	Mouse mRNA for cellu	4.30e-03
7	18	64.3	31720	54	HS1196E3	Human DNA sequence fr	7.35e-01
8	17	60.7	914	43	CHRAABCL	R.americanum chloropl	3.70e+00
9	17	60.7	1147	35	AASPBAB	A.aquillia mRNA for s	3.70e+00
10	17	60.7	2600	37	ZEFEVHIA	Brachydanio rerio (wh	3.70e+00
11	17	60.7	35473	54	HS130G1	Human DNA sequence fr	3.70e+00
12	17	60.7	684973	61	HOMTCRB	Human DNA sequence fr	3.70e+00
13	16	57.1	326	90	HS1AS05SQ	Homo sapiens (clones)	3.70e+00
14	16	57.1	877	90	HOMIFNAM04	Homo sapiens DNA loop	1.76e+01
15	16	57.1	1281	9	HSTP009	Human thyroid peroxid	1.76e+01
16	16	57.1	1281	90	HSTP009	Human thyroid peroxid	1.76e+01
17	16	57.1	2126	52	HSATPNKI	Human Na ⁺ /K ⁺ ATPase	1.76e+01
18	16	57.1	2295	56	HSU33871	Human Down Syndrome	1.76e+01
19	16	57.1	2816	52	HS91K3C	Human DNA sequence fr	1.76e+01
20	16	57.1	3896	9	HS504341	Human Down Syndrome	1.76e+01
21	16	57.1	3896	90	HSU50434	Human Down Syndrome	1.76e+01
22	16	57.1	4208	52	HS91K3D	Human DNA sequence fr	1.76e+01
23	16	57.1	5700	22	PSEBRHA	Pseudomonas sp. bipe	1.76e+01
24	16	57.1	11698	11	KOCTMSCGT	Human 4-hydroxyphenyl	1.76e+01
25	16	57.1	15851	56	HSU29895	K.oxyloca cym A,B,C,D	1.76e+01
26	16	57.1	23065	57	HUMATPSC	Human gene for ATP sy	1.76e+01
27	16	57.1	25719	54	HS1145E5	Human DNA sequence fr	1.76e+01
28	16	57.1	28136	54	HS1184D6	Human DNA sequence fr	1.76e+01
29	16	57.1	32539	54	HS1184D6	Human DNA sequence fr	1.76e+01
30	16	57.1	36171	83	CEC50F4	Caenorhabditis elegans	1.76e+01
31	16	57.1	39756	25	CEC34E11	Caenorhabditis elegans	1.76e+01
32	16	57.1	40592	9	HS1191E1	Human DNA sequence fr	1.76e+01
33	16	57.1	40592	9	HS1191E1	Human DNA sequence fr	1.76e+01
34	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
35	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
36	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
37	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
38	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
39	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
40	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
41	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
42	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
43	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
44	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01
45	16	57.1	41569	90	HS1141A8	Human DNA sequence fr	1.76e+01

ALIGNMENTS

RESULT 1
LOCUS MUSANTP53 536 bp DNA
DEFINITION Mouse tumor antigen p53 gene, 5' end.
ACCESSION M26862
NID g191973
KEYWORDS tumor antigen.
SOURCE Mouse DNA.
ORGANISM Mus musculus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;

REFERENCE 1 Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
AUTHORS 1 (bases 1 to 536)
TITLE Blenz-Houri, R., Libresco, S., Givol, D. and Oren, M.
The 5' region of the p53 gene: Evolutionary conservation and
evidence for a negative regulatory element
JOURNAL EMBO J. 4, 3209-3213 (1985)
MEDLINE 86135950
FEATURES
source location/Qualifiers
1..536
/organism="Mus musculus"
mRNA
321..>536
/note="p53 mRNA"
BASE COUNT 121 a 156 c 120 g 139 t
ORIGIN
Query Match 75.0%; Score 21; DB 65; Length 536;
Best Local Similarity 100.0%; Pred. No. 4.30e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 426 ttcatgtgacatcctgcgt 446
|||||
Oy 8 TTCATTGGACCATCTGCCT 28
RESULT 2
LOCUS MMANT01 536 bp DNA ROD 16-JUN-1995
DEFINITION Murine gene fragment for cellular tumour antigen p53 (exon 1).
ACCESSION X00875
NID g49973
KEYWORDS antigen.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Eukaryota; Eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Myomorpha;
Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 536)
AUTHORS Blenz, R., Zakut-Houri, R., Givol, D. and Oren, M.
TITLE A single gene and a pseudogene for the murine cellular tumour antigen
p53
JOURNAL EMBO J. 3 (9), 2179-2183 (1984)
MEDLINE 85027173
REFERENCE 2 (bases 1 to 536)
AUTHORS Zakut-Houri, R., Oren, M., Blenz, B., Lavie, V., Hazum, S. and Givol, D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306 (5943), 594-597 (1983)
MEDLINE 84068204
FEATURES
source location/Qualifiers
1..536
/organism="Mus musculus"
exon
1..536
/label=ex1
/usedin=X00876:P53_mRNA
replace(157,"a")
conflict /citation=[2]
replace(288,"t")
conflict /citation=[2]
BASE COUNT 122 a 155 c 120 g 139 t
ORIGIN
Query Match 75.0%; Score 21; DB 63; Length 536;
Best Local Similarity 100.0%; Pred. No. 4.30e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 426 ttcatgtgacatcctgcgt 446

Oy 8 TTCATTGGACCATCTGCCT 28
|||||
RESULT 3
LOCUS WMP53G1 554 bp DNA ROD 07-NOV-1985
DEFINITION Mouse genomic DNA for 5' region of cellular tumour antigen p53.
ACCESSION X01235
NID g53572
KEYWORDS antigen.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Eukaryota; Eukaryotes; Metazoa; Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires; Rodentia;
Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 554)
AUTHORS Zakut-Houri, R., Oren, M., Blenz, B., Lavie, V., Hazum, S. and Givol, D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306 (5943), 594-597 (1983)
MEDLINE 84068204
FEATURES
source location/Qualifiers
1..554
/organism="Mus musculus"
misc_feature
1..530
/note="pot. exon"
531..554
intron /note="pot. intron"
BASE COUNT 130 a 154 c 126 g 144 t
ORIGIN
Query Match 75.0%; Score 21; DB 64; Length 554;
Best Local Similarity 100.0%; Pred. No. 4.30e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 420 ttcatgtgacatcctgcgt 440
|||||
Oy 8 TTCATTGGACCATCTGCCT 28
RESULT 4
LOCUS WUSP53A 1285 bp mRNA ROD 05-NOV-1992
DEFINITION Mouse p53 mRNA, complete cds, clone pCD53.
ACCESSION M13872
NID g700198
KEYWORDS p53 gene.
SOURCE Mouse (BALB/c) nontransformed helper T-cell, cDNA to mRNA, clone
pCD-p53.
ORGANISM Mus musculus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE 1 (bases 1 to 1285)
AUTHORS Arai, N., Nomura, D., Yokota, K., Wolf, D., Brill, E., Shohet, O. and
Ritter, V.
TITLE Immunologically distinct p53 molecules generated by alternative
splicing
JOURNAL Mol. Cell. Biol. 6, 3232-3239 (1986)
MEDLINE 87064640
FEATURES
source location/Qualifiers
1..1285
/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"

CDS

/cell_type="nontransformed helper T-cell"
/sequenced_mol="cDNA to mRNA"
113..1285
/codon_start=1
/product="p53 protein"
/db_xref="PID:g200199"
/translation="MTAMEESQSDISLELPISQETFGSLMKLLPPEDILPSPHOMDL
LIPQVVEEFEGPSALVSGAPAADPVTEPEPAAPATPWPPLSSVFSQKTYQG
NYGFHLGFIQSGTAKSWCTSPPLNKLPQIAKTCPPVQLVSAFPAGSVRAMAII
KKSQMTENVVRGCPHHERGSDGGLAPPQHLIRVEGNILPEYLEDROTFRHSVVPYE
PEPAGSEYTTIHKTYMCNCSGCMGMRNPILITITLEDSSGNLIGRDSFEVRVACAPG
RDRRTREENFRKKEVLPPELPGSAKRALPTCTASPPQKKRPIDGEYFTLTKIRGKR
FEMFRELNEALEIKDAHATEESGDSRAHSSYLTKTKGQSTSRHKKTWVKYGPDSU"

BASE COUNT 299 a 385 c 334 g 267 t
ORIGIN

Query Match 75.0%; Score 21; DB 67; Length 1285;
Best Local Similarity 100.0%; Pred. No. 4.30e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 ttcatggagaccatcctgcgt 22
|||||
Oy 8 TTCAATTGGAGACCATCTGCCT 28

RESULT

LOCUS MWP53M 1772 bp mRNA ROD 03-MAY-1985
DEFINITION Mouse p53 cellular tumor antigen, mRNA.
ACCESSION K01700
NID g200204

KEYWORDS DNA-binding protein; antigen; p53 gene; tumor antigen.
SOURCE Mouse embryo F9 carcinoma cells, cDNA to mRNA, clones p176, p211,
p422, p208 [1], and clone pp53-1 [2].

ORGANISM

Mus musculus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.

REFERENCE 1 (bases 1 to 1716)
Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Givol, D.
A single gene and a pseudogene for the cellular tumour antigen p53
Nature 306, 594-597 (1983)

REFERENCE 2 (bases 105 to 1772)
Pennica, D., Goeddel, D.V., Hayflick, J.S., Reich, N.C., Anderson, C.W.
and Levine, A.J.
The amino acid sequence of murine p53 determined from a cDNA clone
Virology 134, 477-482 (1984)

TITLE MEDLINE
JOURNAL MEDLINE
COMMENT 86072076
The murine p53 protein is similar to the avian and human myc gene
products and the adenovirus E1a proteins. All three of these
proteins are rich in proline, contain proline runs or clusters, and
are localized in the nucleus of transformed cells. The E1a proteins
and p53 have been shown to have very short half-lives [2]. The
sequence of cDNA in [1] was established by analysing four separate
cDNA clones; p176 (bp 1-1247), p271 (bp 69-638), p422 (bp
1288-1719), and p208 (bp 1432-1719) [1].
There is only one functional p53 gene in the mouse genome. All
existing different forms of murine p53 must be products of the same
gene, mostly due to post-translational modifications [1]. The
sequence of the murine p53 pseudogene, also reported by [1] (see
separate entry), and the cDNA sequence are almost identical from
nucleotide 186 onward. Upstream of this position the two sequences
deviate totally and no homology can be observed; downstream of here
the two sequences differ by only 4%. The differences are due to
substitutions and to some small deletions or additions in this gene

FEATURES

source

relative to the cDNA [1].
Location/Qualifiers
1..1772

/organism="Mus musculus"
/cell_line="F9"
/dev_stage="embryo"
/sequenced_mol="cDNA to mRNA"
/issue_type="carcinoma"
<1..166
/gene="p53"
/map="11"
/map="11"
158..1330
/gene="p53"
/map="11"
/map="11"

CDS

/codon_start=1
/product="cellular tumor antigen"
/db_xref="PID:g200205"
/translation="MTAMEESQSDISLELPISQETFGSLMKLLPPEDILPSPHOMDL
LIPQVVEEFEGPSALVSGAPAADPVTEPEPAAPATPWPPLSSVFSQKTYQG
NYGFHLGFIQSGTAKSWCTSPPLNKLPQIAKTCPPVQLVSAFPAGSVRAMAII
KKSQMTENVVRGCPHHERGSDGGLAPPQHLIRVEGNILPEYLEDROTFRHSVVPYE
PEPAGSEYTTIHKTYMCNCSGCMGMRNPILITITLEDSSGNLIGRDSFEVRVACAPG
RDRRTREENFRKKEVLPPELPGSAKRALPTCTASPPQKKRPIDGEYFTLTKIRGKR
FEMFRELNEALEIKDAHATEESGDSRAHSSYLTKTKGQSTSRHKKTWVKYGPDSU"

BASE COUNT 385 a 546 c 412 g 429 t
ORIGIN 165 bp upstream of NcoI site.

Query Match 75.0%; Score 21; DB 67; Length 1772;
Best Local Similarity 100.0%; Pred. No. 4.30e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 47 ttcatggagaccatcctgcgt 67
|||||
Oy 8 TTCAATTGGAGACCATCTGCCT 28

RESULT

LOCUS MWP53R 1773 bp RNA ROD 12-SEP-1993
DEFINITION Mouse mRNA for cellular tumour antigen p53.
ACCESSION X01237 K01700
NID g53575

KEYWORDS antigen; tumor antigen.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Eumetazoa; Bilateria; Chordata; Vertebrata; Metazoa; Eumetazoa; Bilateria; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
1 (bases 1 to 1716)
Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Givol, D.
A single gene and a pseudogene for the cellular tumour antigen p53
Nature 306 (5943), 594-597 (1983)

REFERENCE 2 (bases 1 to 1773)
Bienez, B., Zakut-Houri, R., Givol, D. and Oren, M.
Analysis of the gene coding for the murine cellular tumour antigen
p53
EMBO J. 3 (9), 2179-2183 (1984)

JOURNAL MEDLINE
MEDLINE 85027173
REFERENCE 3 (bases 1 to 1773)
Givol, R.
Direct Submission
Submitted (28-NOV-1985) to the EMBL/GenBank/DBJ database

FEATURES

source

Location/Qualifiers

1..1773

/organism="Mus musculus"

158..1330

/note="p53 polypeptide (aa 1-390)"

/codon_start=1

/db_xref="PID:p53576"

/translation="MTAMESQSDISLELPISQETSGLKLIPEDDILPSFHOMDML
LIPQDVEEFEGPSALRYSGAPADPVTETPPVAPAPAPAPPLSSFVSQKTYQG
NYGFHLGFIASGTAKSMQCTSPPLINKLPQVLYTCPPVQVMTAPASRYRAMATY
KKSQHMTEVVRRCRHHKSGSDGGLAPQHLIRVBNLYPEYLEDNQTRHSVVPYE
PPDASGEYTTIHKYMONSCMGMMRPILTIITLEDSSNLGDRSEFVAVACPG
RDRTTEENFRKEVLCPELPESAKRALPTCTASPPQKKRDLDEGYFTLKIRGRR
FEMREINALEIKDAHATEESGDSRAHSYIKTKGQSTSRHKKTWKKVQPPSD"

old_sequence

/note="G is A in [1]"

old_sequence

/citation=[1]

old_sequence

/note="U is C in [1]"

old_sequence

/citation=[1]

old_sequence

/note="U is C in [1]"

old_sequence

/note="U is C in [1]"

old_sequence

/note="A is G in [1]"

old_sequence

/citation=[1]

old_sequence

/note="G is U in [1]"

old_sequence

/citation=[1]

old_sequence

/note="UA is CU in [1]"

old_sequence

/citation=[1]

old_sequence

/note="CU is UC in [1]"

old_sequence

/citation=[1]

old_sequence

/note="U is A in [1]"

old_sequence

/citation=[1]

old_sequence

/note="CU is UC in [1]"

old_sequence

/citation=[1]

old_sequence

/note="G is U in [1]"

old_sequence

/citation=[1]

old_sequence

/note="UA is CU in [1]"

old_sequence

/citation=[1]

old_sequence

/note="U is A in [1]"

old_sequence

/citation=[1]

old_sequence

/note="CU is UC in [1]"

old_sequence

/citation=[1]

old_sequence

/note="G is U in [1]"

old_sequence

/citation=[1]

old_sequence

/note="UA is CU in [1]"

ORGANISM

Homo sapiens

Eukaryota; Mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 31720)

/note="p53 polypeptide (aa 1-390)"

/codon_start=1

/db_xref="PID:p53576"

/translation="MTAMESQSDISLELPISQETSGLKLIPEDDILPSFHOMDML
LIPQDVEEFEGPSALRYSGAPADPVTETPPVAPAPAPAPPLSSFVSQKTYQG
NYGFHLGFIASGTAKSMQCTSPPLINKLPQVLYTCPPVQVMTAPASRYRAMATY
KKSQHMTEVVRRCRHHKSGSDGGLAPQHLIRVBNLYPEYLEDNQTRHSVVPYE
PPDASGEYTTIHKYMONSCMGMMRPILTIITLEDSSNLGDRSEFVAVACPG
RDRTTEENFRKEVLCPELPESAKRALPTCTASPPQKKRDLDEGYFTLKIRGRR
FEMREINALEIKDAHATEESGDSRAHSYIKTKGQSTSRHKKTWKKVQPPSD"

/citation=[1]

/note="G is A in [1]"

/note="U is C in [1]"

/note="U is C in [1]"

/note="A is G in [1]"

/note="U is C in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

REFERENCE

1 (bases 1 to 31720)

/note="p53 polypeptide (aa 1-390)"

/codon_start=1

/db_xref="PID:p53576"

/translation="MTAMESQSDISLELPISQETSGLKLIPEDDILPSFHOMDML
LIPQDVEEFEGPSALRYSGAPADPVTETPPVAPAPAPAPPLSSFVSQKTYQG
NYGFHLGFIASGTAKSMQCTSPPLINKLPQVLYTCPPVQVMTAPASRYRAMATY
KKSQHMTEVVRRCRHHKSGSDGGLAPQHLIRVBNLYPEYLEDNQTRHSVVPYE
PPDASGEYTTIHKYMONSCMGMMRPILTIITLEDSSNLGDRSEFVAVACPG
RDRTTEENFRKEVLCPELPESAKRALPTCTASPPQKKRDLDEGYFTLKIRGRR
FEMREINALEIKDAHATEESGDSRAHSYIKTKGQSTSRHKKTWKKVQPPSD"

/citation=[1]

/note="G is A in [1]"

/note="U is C in [1]"

/note="U is C in [1]"

/note="A is G in [1]"

/note="U is C in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

/note="U is A in [1]"

/note="CU is UC in [1]"

/note="G is U in [1]"

/note="UA is CU in [1]"

/note="CU is UC in [1]"

```
misc_feature      /note="Alu repeat: matches 308..1 of consensus"  
4985..5650  
/note="match: multiple ESTs"  
repeat_region    5744..5794  
/note="17 copies of 3 mer 80 % conserved"  
6380..6409  
/partial  
repeat_region    /note="Alu repeat: matches 1..31 of consensus"  
6411..6582  
/partial  
repeat_region    /note="Alu repeat: matches 73..259 of consensus"  
8172..8342  
/partial  
repeat_region    /note="Alu repeat: matches 308..124 of consensus"  
9498..9591  
/partial  
repeat_region    /note="Alu repeat: matches 308..214 of consensus"  
9592..9748  
/partial  
repeat_region    /note="Alu repeat: matches 175..1 of consensus"  
10969..11257  
/note="Alu repeat: matches 1..308 of consensus"  
11814..11858  
/note="3 copies of 15 mer 93 % conserved"  
12218..12415  
/partial  
repeat_region    /note="Alu repeat: matches 1..214 of consensus"  
12416..12478  
/partial  
repeat_region    /note="Alu repeat: matches 248..308 of consensus"  
12498..12613  
/partial  
repeat_region    /note="Alu repeat: matches 118..1 of consensus"  
13089..13418  
/note="THE18 element fragment"  
13742..13846  
/partial  
repeat_region    /note="Alu repeat: matches 148..29 of consensus"  
14100..14303  
/partial  
misc_feature     /note="Alu repeat: matches 85..301 of consensus"  
complement(14579..16077)  
/note="match: EST T03845 clone gt43"  
15104..15166  
/note="MIR element fragment"  
18047..18146  
/note="MIR element fragment"  
19465..19782  
/note="match: EST T03956 clone gt133"  
21939..22228  
/note="Alu repeat: matches 1..308 of consensus"  
22381..22555  
/note="L1 element fragment"  
23074..23108  
/partial  
repeat_region    /note="Alu repeat: matches 157..191 of consensus"  
23109..23205  
/partial  
repeat_region    /note="Alu repeat: matches 42..143 of consensus"  
23258..23538  
/note="Alu repeat: matches 11..308 of consensus"  
23768..24071  
/note="Alu repeat: matches 1..308 of consensus"  
24223..24306
```

```
repeat_region    /note="L1 element fragment"  
24378..24527  
/partial  
repeat_region    /note="Alu repeat: matches 152..304 of consensus"  
24590..24881  
/note="Alu repeat: matches 1..308 of consensus"  
24895..25202  
/note="Alu repeat: matches 1..308 of consensus"  
25205..25361  
/note="L1 element fragment"  
25371..25663  
/note="Alu repeat: matches 1..308 of consensus"  
25853..26140  
/note="Alu repeat: matches 1..308 of consensus"  
26141..26314  
/note="L1 element fragment"  
26392..26683  
/note="L1 element fragment"  
26728..26862  
/note="L1 element fragment"  
26869..27160  
/note="Alu repeat: matches 1..308 of consensus"  
27248..27407  
/partial  
repeat_region    /note="Alu repeat: matches 1..177 of consensus"  
27408..27523  
/partial  
repeat_region    /note="Alu repeat: matches 190..308 of consensus"  
27610..27730  
/note="L1 element fragment"  
27780..28072  
/note="Alu repeat: matches 308..1 of consensus"  
28089..28134  
/note="23 copies of 2 mer 83 % conserved"  
28142..28398  
/note="L1 element fragment"  
28459..28557  
/note="L1 element fragment"  
28581..28869  
/note="L1 element fragment"  
28873..29722  
/note="Alu repeat: matches 308..1 of consensus"  
29782..30151  
/note="L1 element fragment"  
30243..30521  
/note="L1 element fragment"  
30679..30805  
/partial  
/note="Alu repeat: matches 144..1 of consensus"
```

```
BASE COUNT      9652 a   7283 c   7157 g   7628 t  
ORIGIN
```

```
Query Match      64.3%; Score 18; DB 54; Length 31720;  
Best Local Similarity 100.0%; P-Id No. 7.35e-01;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 27323 atggaccatcctcgct 27340
```

```
0y 11 ATTGGGACCATCTCGCT 28
```

```
RESULT 8          CHRABCL 914 bp DNA          PLN          11-FEB-1994  
LOCUS  
DEFINITION R.americanum chloroplast rbcL gene for
```

ACCESSION ribulose-1,5-bisphosphate-carboxylase and promoter region.
X69751
NID g12204
KEYWORDS atpB gene; carbon dioxide fixation; photosynthesis;
promoter region; rbcL gene.
SOURCE Ribes americanum.
ORGANISM Chloroplast Ribes americanum
Eukaryotae; mitochondrial eukaryotes; Chlorophyta/Embryophyta
group; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
Magnoliopsida; Rosidae; Rosales; Grossulariaceae; Ribes.
REFERENCE 1 (bases 1 to 914)
AUTHORS Savolainen, V.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-1992) to the EMBL/Genbank/DBJ databases. V.
Savolainen, Conservatoire et Jardin Botaniques, 1 chemin de
l'Imperatrice, 1292 Chambesey, Geneva, SWITZERLAND
REFERENCE 2 (bases 1 to 914)
AUTHORS Savolainen, V., Manen, J.F., Douzery, E. and Spichiger, R.
TITLE Molecular phylogeny of families related to Celastrales based on
rbcL 5' flanking sequences
JOURNAL Mol. Phylogenet. Evol. 3 (1), 27-37 (1994)
MEDLINE 94297805
FEATURES
source location/Qualifiers
1..914
/organism="Ribes americanum"
/chloroplast
/isolate="sample RA12"
753..>914
/gene="rbcL"
/codon_start=1
/product="ribulose-1,5-bisphosphate-carboxylase"
/db_xref="PID:g12205"
/translation="MSPQETKASVGRKAGKVDKLTYYTEYKDTLLAFRYTP
QPCVPEENG"
BASE COUNT 306 a 148 c 149 g 311 t
ORIGIN
Query Match 60.7%; Score 17; DB 43; Length 914;
Best Local Similarity 90.5%; Pred.No. 3.70e+00;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 101 tcgattcttgaggaccatc 121
||||| ||| ||||| ||||| |||
Qy 3 TCGATTCTATTGGACCATCC 23
RESULT 9
LOCUS AASPAB 1147 bp RNA VRT 01-JUN-1994
DEFINITION A.anguilla mRNA for sodium/potassium ATPase, beta subunit.
ACCESSION X76109
NID g496529
KEYWORDS sodium/potassium ATPase.
SOURCE European eel.
ORGANISM Anguilla anguilla
Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Actinopterygii;
Actinopteri; Holostei; Halecoetomi; Teleostei; Elopocephala;
Elopomorphia; Anguilliformes; Anguillioidei; Anguillidae; Anguilla.
REFERENCE 1 (bases 1 to 1147)
AUTHORS Cutler, C., Sanders, I.L. and Cramb, G.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1147)
AUTHORS Cutler, C.P.

FEATURES
source location/Qualifiers
1..1147
/organism="Anguilla anguilla"
/isolate="Am1.1, Anb2.13"
/tissue_type="gill"
97..1008
/gene="ATNB_AGCAN"
/EC_number="3.6.1.37"
/codon_start=1
/product="sodium /potassium-transporting ATPase, beta
subunit"
/db_xref="PID:g496530"
/translation="MPATKDSGCKKFTLMSKKKELGRTGSGMAKILFVYFYG
CLAGLFGTIGALLTINDFRPYQDRVAPPGSHTPSKSESMKXFDPSYOKYV
KAMHDFLQAYNDSKQENMKEDCDTPKSYINRGELNNQGIKXKCLFRRSMIDKX
GLEDPTEGSESKPCLIYKILNRIVNRRPRPTSNISPEEAQSKYQDPVPIYCTNKR
EEDAKVREIKRYGIQEGEPDQITPYGKQHPQIDP LVAVHFTNLMTATELRICR
VYGMNIAYSDDRYRGRFDVKTFINES"
1116..>1147
misc_feature
BASE COUNT 338 a 277 c 261 g 271 t
ORIGIN
Query Match 60.7%; Score 17; DB 35; Length 1147;
Best Local Similarity 90.5%; Pred.No. 3.70e+00;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 243 attcattggcaccatctgc 27
||||| ||||| ||||| |||
Qy 7 ATTCATTGGACCATCTCGC 27
RESULT 10
LOCUS ZEFVNHIA 2600 bp mRNA VRT 25-AUG-1994
DEFINITION Brachydanio rerio (vhh-1) mRNA, complete cds.
ACCESSION U27585
NID g452159
KEYWORDS floor plate induction; motor neuron induction;
vertebrate hedgehog homologue; vhh-1 gene.
SOURCE zebrafish cDNA to mRNA.
ORGANISM Brachydanio rerio
Eukaryotae; Animalia; Chordata; Vertebrata; Osteichthyes;
Actinopterygii; Cypriniformes; Cyprinoidae; Cyprinidae.
REFERENCE 1 (bases 1 to 2600)
AUTHORS Roeling, H., Augsburg, A., Heemkerk, J., Korzh, V., Norlin, S., Ruiz
I Altaba, A., Tanabe, Y., Placzek, M., Edlund, T., Jessell, T.M. and
Dodd, J.
TITLE Floor plate and motor neuron induction by vhh-1, a vertebrate
homolog of hedgehog expressed by the notochord
JOURNAL Cell 76 (4), 761-775 (1994)
MEDLINE 94170375
FEATURES
source location/Qualifiers
1..2600
/organism="Brachydanio rerio"
/dev_stage="9-16 hours post fertilization"
/sequenced_mol="cDNA to mRNA"
/tissue_type="total embryo"
/tissue_lib="lambda ZAP-II"
320..388
/gene="vhh-1"

```
CDS
/codon_start=1
320..1576
/feature="vhh-1"
/standard_name="vertebrate hedgehog homolog"
/codon_start=1
/db_xref="PID:g452160"
/translation="MRLRLVLLVSLTLISLVSGLAGCGRGYGRRRHKKLPLAY
KQIPVNAEKTLAGSRYEKQITNSERKEITPNIPDIIFKDENTGADRLMTOHC
KQKLSIAISYNNHMPGVKLRVTEGMDHGFEEIHYEGRAVDITSDRKSRYGT
LSRLAVEGPDWVYSEKAIHGSVAENSVAAKSGGCPGSAIVSLDQGGKAVDGL
NPGDKVLAADSAKLVFSDFIMFTDDSTTRRVTYLETQEPVEKTIITAAHLEVLID
NSTEDLHTTAAYASVRAQGVWVDSDQAKSVIYORIYEEORGFAPVTAGTIT
VVDRLASCVAIVEDOGLAHAFAPARLALYYVSSEFLPQNSSSNATIQDQGVHWYS
RLLYQMGTYLDDSNMHP LGMSVNS"
mat_peptide
389..1573
/feature="vhh-1"
/standard_name="vertebrate hedgehog homolog"
/codon_start=1
BASE COUNT      715 a      581 c      661 g      643 t
ORIGIN
```

```
Query Match      60.7%; Score 17; DB 37; Length 2600;
Best Local Similarity 87.0%; Pred. No. 3.70e+00;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Db 1809 agtcgaatgcagtgtacatcc 1831
||||||| || || |||||
0y 1 ACTGCAATTCATTGGACCATCC 23
```

```
RESULT 11
LOCUS      HSL30G1      35473 bp      DNA      PRI      23-JAN-1996
DEFINITION Human DNA sequence from cosmid L30G1, Huntington's Disease Region,
              chromosome 4p16.3.
ACCESSION  Z68870
NID        q1164910
KEYWORDS   4p16.3.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 35473)
AUTHORS    Heathcote, R.
TITLE      Direct Submission
JOURNAL    Submitted (22-JAN-1996) Sanger Centre, Hinxton, Cambridgeshire,
              CB10 1RQ, England. E-mail enquires: humquery@sanger.ac.uk
COMMENT    IMPORTANT:
            This sequence is the entire insert of clone L30G1. This sequence
            has been finished according to sequence map criteria as follows.
            An attempt is made to resolve all sequencing problems, such as
            compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            The true left end of clone L30G1 is at 1 in this sequence. The true
            right end of clone L26E12 is at 21857.
            The true right end of clone L30G1 is at 35473.
            L30G1 is from cosmid library LA04NC01 constructed at the Human
            Genome Center, Los Alamos National Laboratory, NM 87545 under the
            auspices of the U.S. Department of Energy. The library was
            constructed using flow-sorted human chromosome 4 from a
            Haesler-human hybrid cell line (UV20HL21-27) containing human
            chromosomes 4, 8 and 21.
            VECTOR: sCos1
```

L30G1 is contained in a clone contig spanning
2Mb which is described in Baxendale et al, Nature Genetics 4 (1993
) 181-186. See also Myers et al, Cytogenet Cell Genet. 66 (1994
) 218-230.

```
FEATURES
source
1..35473
location/qualifiers
/organism="Homo sapiens"
/chromosome="4"
/map="4p16.3"
/clone="L30G1"
/clone_1lb="LA04NC01"
/cell_line="UV20HL21-27"
53..362
/misc_feature
/feature="match: 5' EST T68903 clone 82158"
385..629
/feature="match: 5' EST R50568 clone 153110"
731..1035
/feature="match: multiple ESTs"
1172..1600
/feature="MT2A1 element fragment"
1779..2072
/partial
/feature="Alu repeat: matches 308..1 of consensus"
2532..2824
/partial
/feature="Alu repeat: matches 308..1 of consensus"
4331..4708
/feature="match: multiple ESTs"
4618..4697
/feature="2 copies of 40 mer 95 % conserved"
5085..5382
/partial
/feature="Alu repeat: matches 308..1 of consensus"
5393..5637
/partial
/feature="Alu repeat: matches 308..44 of consensus"
5638..6154
/feature="L1 element fragment"
6265..6372
/feature="L1 element fragment"
6361..6442
/feature="MSTC element fragment"
6533..6732
/feature="MSTC element fragment"
6661..6732
/feature="MSTC element fragment"
7325..7605
/feature="Alu repeat: matches 1..297 of consensus"
10115..10512
/feature="match: multiple ESTs"
11204..11314
/partial
/feature="Alu repeat: matches 1..112 of consensus"
11314..11483
/partial
/feature="Alu repeat: matches 308..132 of consensus"
11484..11775
/partial
/feature="Alu repeat: matches 308..1 of consensus"
11793..11964
/partial
/feature="Alu repeat: matches 308..117 of consensus"
11991..12281
/partial
```

```
/note="Alu repeat: matches 308..1 of consensus"
repeat_region 12484..12776
/note="Alu repeat: matches 1..308 of consensus"
repeat_region 12779..13009
/partial
/note="Alu repeat: matches 1..246 of consensus"
repeat_region 13708..13779
/partial
/note="Alu repeat: matches 245..308 of consensus"
repeat_region 14323..14382
/partial
/note="Alu repeat: matches 1..61 of consensus"
repeat_region 14388..14451
/partial
/note="Alu repeat: matches 81..148 of consensus"
repeat_region 14458..14740
/note="Alu repeat: matches 1..308 of consensus"
repeat_region 14766..14802
/partial
/note="Alu repeat: matches 145..185 of consensus"
repeat_region 15249..15288
/partial
/note="Alu repeat: matches 270..308 of consensus"
repeat_region 15428..15713
/partial
/note="Alu repeat: matches 308..1 of consensus"
repeat_region 15733..16045
/partial
/note="Alu repeat: matches 308..1 of consensus"
repeat_region 16356..16494
/partial
/note="Alu repeat: matches 1..155 of consensus"
repeat_region 16496..16787
/note="Alu repeat: matches 1..308 of consensus"
repeat_region 16794..16975
/partial
/note="Alu repeat: matches 123..308 of consensus"
repeat_region 17235..17441
/partial
/note="Alu repeat: matches 308..88 of consensus"
repeat_region 17449..17484
/note="18 copies of 2 mer 83 & conserved"
repeat_region 17516..17806
/note="Alu repeat: matches 1..306 of consensus"
repeat_region 17975..18042
/partial
/note="Alu repeat: matches 308..240 of consensus"
repeat_region 18045..18125
/partial
/note="Alu repeat: matches 83..1 of consensus"
repeat_region 18736..18874
/partial
/note="Alu repeat: matches 159..308 of consensus"
repeat_region 19106..19392
/partial
/note="Alu repeat: matches 308..1 of consensus"
repeat_region 19406..19697
/partial
/note="Alu repeat: matches 306..1 of consensus"
repeat_region 19733..19843
/partial
/note="Alu repeat: matches 1..118 of consensus"
repeat_region 21194..21362
/partial
```

```
/note="Alu repeat: matches 308..135 of consensus"
repeat_region 21363..21653
/partial
/note="Alu repeat: matches 308..1 of consensus"
repeat_region 21654..21793
/partial
/note="Alu repeat: matches 161..1 of consensus"
repeat_region 22026..22308
/partial
/note="Alu repeat: matches 308..1 of consensus"
repeat_region 22591..22875
/note="Alu repeat: matches 1..296 of consensus"
repeat_region 22934..23220
/partial
/note="Alu repeat: matches 306..1 of consensus"
repeat_region 23365..23482
/partial
/note="Alu repeat: matches 308..190 of consensus"
repeat_region 23504..23617
/note="L1 element fragment"
repeat_region 23952..24239
/partial
/note="Alu repeat: matches 308..1 of consensus"
repeat_region 24730..24908
/partial
/note="Alu repeat: matches 1..193 of consensus"
repeat_region 24919..25064
/partial
/note="Alu repeat: matches 308..164 of consensus"
repeat_region 25079..25373
/partial
/note="Alu repeat: matches 308..1 of consensus"
repeat_region 25421..25692
/partial
/note="Alu repeat: matches 17..308 of consensus"
repeat_region 25709..25999
/note="Alu repeat: matches 1..308 of consensus"
misc_feature complement(26350..26616)
/partial

...
Note: remainder of annotations omitted.

Query Match 60.7%; Score 17; DB 54; Length 35473;
Best Local Similarity 94.7%; Pred. No. 3.70e+00;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 32887 agccctgagtgtcccaatg 32905
|||||
Cp 28 AGCCAGATGCTCCCAATG 10

RESULT 12
LOCUS HUMHCRB 684973 bp DNA PRI 14-JAN-1995
DEFINITION Homo sapiens (clones: K41A, K35, K26, K56, X21B, G54, H137, H18,
H18/G15gap, G15, X1A, A27, A212, A14, H7.1, H12.18, H130, A16,
C215, G1, C68, C21, X11, X6A, CBG1, CBG1/C29gap, and C29) germline
T-cell receptor beta chain, complete gene.
ACCESSION I36092
MID g540583
KEYWORDS C-beta gene segment; C-region; D-region; J-beta gene segment;
J-segment; T cell receptor beta chain; TCR-beta gene;
V-beta gene segment; V-segment; cell membrane protein;
constant region; diversity region; germline; joining segment;
trypsin; trypsinogen; variable segment.
```



```
repeat_unit      /rpt_family="Alu"  
                  /note="putative"  
                  /rpt_family="MT1A"  
repeat_unit      complement(78720..78991)  
                  /note="putative"  
                  /rpt_family="Alu"  
repeat_unit      complement(79236..79567)  
                  /note="putative"  
                  /rpt_family="MST1"  
mRNA             complement(join(<79799..79948,80489..80625,81147..81400,  
82186..82348,84121..>84159))  
                  /gene="TRYF"  
exon             complement(<79799..79948)  
                  /gene="TRYF"  
                  /note="putative"  
                  /pseudo  
                  /number=5  
CDS              complement(join(79799..79948,80489..80625,81147..81400,  
82186..82348,84121..84159))  
                  /gene="TRYF"  
                  /note="putative"  
                  /pseudo  
                  /codon_start=1  
intron           complement(79949..80488)  
                  /gene="TRYF"  
                  /note="putative"  
                  /number=4  
exon             complement(80489..80625)  
                  /gene="TRYF"  
                  /note="putative"  
                  /pseudo  
                  /number=4  
intron           complement(80626..81146)  
                  /gene="TRYF"  
                  /note="putative; does not fit consensus"  
                  /number=3  
exon             complement(81147..81400)  
                  /gene="TRYF"  
                  /note="putative"  
                  /pseudo  
                  /number=3  
intron           complement(81401..82185)  
                  /gene="TRYF"  
                  /note="putative"  
                  /number=2  
exon             complement(82186..82348)  
                  /gene="TRYF"  
                  /note="putative"  
                  /pseudo  
                  /number=2  
intron           complement(82349..84120)  
                  /gene="TRYF"  
                  /note="putative"  
                  /number=1  
                  /note="putative"  
                  /rpt_family="Alu"  
                  complement(84121..>84159)  
                  /gene="TRYF"  
exon
```

..
Note: remainder of annotations omitted.

Query Match 60.7%; Score 17; DB 61; Length 684973;

```
Best Local Similarity 100.0%; Pred. No. 3.70e+00;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 523902 atggagaccatcctgcg 523918  
Oy 11 ATTGGAGCACCATCTCGCT 27  
|||||
```

```
RESULT 13  
LOCUS HSLAS05SQ 326 bp DNA PRI 22-APR-1996  
DEFINITION H.sapiens DNA loop attachment sequence (clone LAS05).  
ACCESSION X91841  
NID g1009421  
KEYWORDS Alu sequence; loop attachment sequence.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Eukaryota; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
```

```
REFERENCE 1 (bases 1 to 326)  
AUTHORS Cook, P. R.  
TITLE Direct Submission  
JOURNAL Submitted (25-SEP-1995) P. R. Cook, Sir William Dunn School of  
Pathology, University of Oxford, South Parks Road, Oxford, OX1 3RE,  
UK
```

```
REFERENCE 2 (bases 1 to 326)  
AUTHORS Jackson, D.A., Bartlett, J.M., and Cook, P. R.  
TITLE Sequence attaching loops of nuclear and mitochondrial DNA to  
underlying structures in human cells: the role of transcription  
units  
JOURNAL Nucleic Acids Res. 24, 1212-1219 (1996)  
COMMENT Related sequence:  
J.Jurka, Proc.Natl.Acad.Sci.USA, 85, 4775-4778, 1988.
```

```
FEATURES  
source NCBI gi: 1009421  
location/Qualifiers  
1..326  
/organism="Homo sapiens"  
/cell_line="HeLa"  
/clone_lib="DNA loop attachment sequences (LAS)"  
/clone="LAS05"  
/map="48 nucleotide number"  
repeat_region 1..246  
/rpt_family="Alu"
```

```
misc_feature <1..>326  
/note="DNA loop attachment sequences"  
BASE COUNT 118 a 72 c 80 g 56 t  
ORIGIN
```

```
Query Match 57.1%; Score 16; DB 90; Length 326;  
Best Local Similarity 94.4%; Pred. No. 1.76e+01;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Db 29 atggagaccatcctgcgct 46  
Oy 11 ATTGGAGCACCATCTCGCT 28  
|||||
```

```
RESULT 14  
ID HSLAS05SQ standard; DNA; PRI; 326 BP.  
AC X91841;  
DT 05-OCT-1995 (Rel. 45, Created)  
DT 22-APR-1996 (Rel. 47, Last updated, Version 5)  
DE H.sapiens DNA loop attachment sequence (clone LAS05)  
KW Alu sequence; loop attachment sequence.
```


OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
RN [2]
RP 1-326
RA Cook P.R.;
RT ;
RL Submitted (25-SEP-1995) to the EMBL/GenBank/DBJ databases.
RL P.R. Cook, Sir William Dunn School of Pathology, University of
RL Oxford, South Parks Road, Oxford, OX1 3RE, UK
RN [3]
RP 1-326
RA Jackson D.A., Bartlett J.M., Cook P.R.;
RT *Sequence attaching loops of nuclear and mitochondrial DNA to
RT underlying structures in human cells: the role of transcription
RT units*;
RL Nucleic Acids Res. 24:1212-1219(1996).
CC Related sequence:
CC J. Jurka, Proc.Natl.Acad.Sci.USA, 85, 4775-4778, 1988
FH Key Location/Qualifiers
FH 1..326
FT source /organism="Homo sapiens"
FT /cell_line="HeLa"
FT /clone_lib="DNA loop attachment sequences (LAS)"
FT /clone="LAS05"
FT /map="48 nucleotide number"
FT misc_feature <1..>326
FT /note="DNA loop attachment sequences"
FT repeat_region 1..246
FT /rpt_family="Alu"
SQ Sequence 326 BP; 118 A; 72 C; 80 G; 56 T; 0 other;

Query Match 57.1%; Score 16; DB 9; Length 326;
Best Local Similarity 94.4%; Pred. No. 1.76e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 29 attgagacatccctgct 46
||| |||||||||
Gy 11 ATTGGGACATCTGCTGCT 28

RESULT 15
LOCUS HUMIFNM04 877 bp DNA PRI 04-APR-1996
DEFINITION Homo sapiens (clone Q-20D3) interferon receptor (IFNAR2) gene, exon
5.
ACCESSION LA2240
NTD g994719
KEYWORDS cytokine receptor; helical cytokine receptor family;
SEGMENT hematopoietin receptor family; interferon receptor.
SOURCE 4 of 8
ORGANISM Homo sapiens (clone: Q-20D3) (clone library: LT21NC02-Q) DNA.
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 877)
AUTHORS Lutfalla,G., Holland,S.J., Cinato,E., Monneron,D., Reboul,J.,
Rogers,N.C., Smith,J.M., Stark,G.R., Gardiner,K., Mogensen,K.E.,
Kerr,I.M. and Uze,G.
TITLE Mutant 5'3A cells are complemented by an interferon-alpha beta
receptor subunit generated by alternative processing of a new
member of a cytokine receptor gene cluster
JOURNAL EMBO J. 14 (20), 5100-5108 (1995)
MEDLINE 96067138

COMMENT NCBI gi: 994719
FEATURES Location/Qualifiers
source 1..877
/organism="Homo sapiens"
/macronuclear
/clone="Q-20D3"
/clone_lib="LT21NC02-Q"
/map="21q22.1"
order(L42239;1569..1661,1..581)
/partial
/note="1.6 kb"
/number=4
exon 582..754
/gene="IFNAR2"
/note="encodes SD100A s3-s7"
/number=5
exon 582..754
/gene="IFNAR2"
/note="encodes SD100A s3-s7"
/number=5
BASE COUNT 237 a 212 c 200 g 228 t
ORIGIN

Query Match 57.1%; Score 16; DB 90; Length 877;
Best Local Similarity 94.4%; Pred. No. 1.76e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 145 acccagatgctcatt 162
|||||
Cp 28 AGCCAGATGCTCCCAAT 11

Search completed: Tue Mar 25 02:47:15 1997
Job time : 63 secs.

Mar 25 02:50

US-08-644-289-4.rng

1

WIRELESS
(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Merch_nm n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Mar 25 02:52:36 1997; Maspar time 9.33 Seconds

280.530 Million cell updates/sec

Tabular output not generated.

Title: >US-08-644-289-4

Description: (1-30) from US08644289.seq

Perfect Score: 30

N.A. Sequence: 1 AGTCGATCTCGAGTCAGCCCTCTGTCT 30

Comp: TCAGCCTAGACCTCACTCGGACGACAGA

Scoring table: TABLE default

Gap 10

Mismatch STD : Dbase 0; Query 0

Searched: 113505 seqs, 43611913 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

n-geneseq25
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22

Statistics: Mean 5.779; Variance 3.522; scale 1.641

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description	Pred. No.
1	24	80.0	91	9	051746	6.72e-04
2	21	70.0	1173	12	067883	3.08e-02
3	20	66.7	91	9	051746	1.07e-01
4	19	63.3	34	21	T10488	3.63e-01
5	17	56.7	471	11	062363	3.91e+00
6	17	56.7	984	11	062361	3.91e+00
7	17	56.7	1062	11	062359	3.91e+00
8	17	56.7	1173	5	Q29632	3.91e+00

Mar 25 02:50

US-08-644-289-4.rng

2

c	9	17	56.7	1181	15	Q97854	Human p53 cDNA.	3.91e+00
c	10	17	56.7	1182	22	T29719	Wild type p53 gene se	3.91e+00
c	11	17	56.7	1182	12	067884	Human p53 DNA.	3.91e+00
c	12	17	56.7	1182	3	Q22995	Sequence encoding 53	3.91e+00
c	13	17	56.7	1483	12	067864	Vaccinia H6 promoter/	3.91e+00
c	14	17	56.7	1512	12	067863	Vaccinia H6 promoter/	3.91e+00
c	15	17	56.7	9502	13	Q74770	Hepatitis C virus gen	3.91e+00
c	16	16	53.3	71	13	Q80398	CDR-grafted L243-gH p	1.24e+01
c	17	16	53.3	71	13	Q80676	L243-gH assembly olig	1.24e+01
c	18	16	53.3	435	13	Q80434	CDR-grafted L243-gH V	1.24e+01
c	19	16	53.3	435	13	Q80363	Humanized antibody L2	1.24e+01
c	20	16	53.3	3017	1	Q05271	Sequence encoding sec	1.24e+01
c	21	15	50.0	274	1	N92102	Sequence of the hepat	3.79e+01
c	22	15	50.0	438	22	T34165	Mouse ob cDNA clone p	3.79e+01
c	23	15	50.0	450	18	T19056	Human gene signature	3.79e+01
c	24	15	50.0	456	1	N92101	Sequence of the hepat	3.79e+01
c	25	15	50.0	456	1	N90331	Hepatitis C virus (HC	3.79e+01
c	26	15	50.0	504	22	T34163	Mouse ob cDNA clone p	3.79e+01
c	27	15	50.0	854	1	Q03296	Sequence of clone c21	3.79e+01
c	28	15	50.0	1173	7	Q43893	NANB hepatitis virus	3.79e+01
c	29	15	50.0	1180	14	Q79144	Hepatitis C virus gen	3.79e+01
c	30	15	50.0	1191	14	Q79142	Hepatitis C virus gen	3.79e+01
c	31	15	50.0	1847	1	N90719	Paes-3 clone encoding	3.79e+01
c	32	15	50.0	2793	21	T16372	Obesity protein codin	3.79e+01
c	33	15	50.0	4337	14	Q82833	Human stromalin-1 DNA	3.79e+01
c	34	15	50.0	4757	17	Q98957	Chicken genomic DNA f	3.79e+01
c	35	15	50.0	6905	1	N92103	Combined open reading	3.79e+01
c	36	15	50.0	7310	1	N90336	Composite hepatitis C	3.79e+01
c	37	15	50.0	7310	20	Q98221	Hepatitis C virus clo	3.79e+01
c	38	15	50.0	7350	17	Q98955	Complete B-G gene enc	3.79e+01
c	39	15	50.0	8316	1	Q05955	Hepatitis C virus ope	3.79e+01
c	40	15	50.0	9185	2	Q10566	Hepatitis C virus str	3.79e+01
c	41	15	50.0	9185	1	Q05956	Sense strand of the c	3.79e+01
c	42	15	50.0	9379	6	Q36209	Composite cDNA for He	3.79e+01
c	43	15	50.0	9400	3	Q21744	Compilied HCV cDNA.	3.79e+01
c	44	15	50.0	9416	3	Q22871	NANBV Huhc c59 isola	3.79e+01
c	45	15	50.0	12001	13	Q76213	HSV L/ST region.	3.79e+01

ALIGNMENTS

RESULT	1
ID	051746 standard; cDNA: 91 BP.
AC	051746/
DE	31-MAY-1994 (first entry)
DT	Oligonucleotide probe MK14-A
KM	Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW	ss.
OS	Synthetic.
PN	EP-571911-A.
PD	01-DEC-1993.
PF	24-MAY-1993; 108325.
PR	26-MAY-1992; US-889651.
PA	(BECT) BECTON DICKINSON CO.
PI	Shank DD, Spears PA;
DR	MP1; 93-378844/48.
PT	New oligo:nucleotide probes specific for Mycobacteria - used for
PT	detection and amplification of Mycobacteria nucleic acid in
PT	samples
PS	Claim 3; Page 14; 23pp; English.
CC	Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC	(Q51735). It hybridized to all spp. of mycobacteria tested, but
CC	cross reacted to a few non-mycobacterial spp. The probe may
CC	be useful as an initial screen for mycobacterial infection.

CC (p53s) and their expression. When the antibodies are injected into
CC cells they may cause cell cycle arrest. Vectors contg. p53s cDNA may
CC be used in gene therapy of cancers and other proliferative disorders
CC e.g. proliase.
SQ Sequence 34 BP; 5 A; 8 C; 12 G; 9 T;

Query Match 63.3%; Score 19; DB 21; Length 34;
Best Local Similarity 100.0%; Pred. No. 3.63e-01;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 agtcgactcctgagtgag 19
|||||
1 ACTCGACTCTCGACTGAG 19

RESULT 5

ID Q62363 standard; cDNA; 471 BP.

AC Q62363;

DT 18-NOV-1994 (first entry)

DE Fragment coding for human p53 amino acids 237-393.

KM Human nuclear phosphoprotein p53; tumour suppressor gene product;

OS anti-oncogene; cancer; tumour; antibody binding region; epitope; ds.

05 Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..471

FT /tag= a

FT /product= p53(237-393)

FT /transl_except= pos:109..111,aa:Arg

FT /note= "partial CDS, therefore sequence does not

FT contain start or stop codons"

FN M09408241-A.

PD 14-APR-1994.

PF 30-SEP-1993; E02666.

PR 30-SEP-1992; DE-232823.

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM STIFTUNG.

PI Klein R, Schranz P, Tessmer C, Volkmann M, Zentgraf H;

DR WPI; 94-135732/16.

DR P-PSDB; R51878.

PT Non-radioactive detection of p53 specific antibodies - by capture

PT on immobilised p53 or its fragments, then reaction with labelled

PT second antibody, for diagnosis of tumours and suitable for

PT screening

PS Claim 13; Page 25; 35pp; German.

CC Antibodies specific for p53 are detected by binding to immobilised

CC fragments of the p53 gene product containing the antibody-binding

CC region. Preferred fragments contain amino acids 1-241, 40-349,

CC 40-393, 66-241, 66-393, 237-349, 237-393 and esp. 9-33, 37-52 or

CC 368-386. See Q62357-Q62366 for nucleic acid sequences coding for

CC each of these fragments.

SQ Sequence 471 BP; 128 A; 133 C; 132 G; 78 T;

Query Match 56.7%; Score 17; DB 11; Length 471;

Best Local Similarity 100.0%; Pred. No. 3.91e+00;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 376 agcaggtcactccag 392
|||||

Cp 26 AGCAGGCTCCTCAG 10

RESULT 6
ID Q62361 standard; cDNA; 984 BP.
AC Q62361;
DT 18-NOV-1994 (first entry)

DE Fragment coding for human p53 amino acids 66-393.
KM Human nuclear phosphoprotein p53; tumour suppressor gene product;
KM anti-oncogene; cancer; tumour; antibody binding region; epitope; ds.
OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..984

FT /tag= a

FT /product= p53(66-393)

FT /transl_except= pos:622..624,aa:Arg

FT /note= "partial CDS, therefore sequence does not

FT contain start or stop codons"

FN M09408241-A.

PD 14-APR-1994.

PF 30-SEP-1993; E02666.

PR 30-SEP-1992; DE-232823.

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM STIFTUNG.

PI Klein R, Schranz P, Tessmer C, Volkmann M, Zentgraf H;

DR WPI; 94-135732/16.

DR P-PSDB; R51876.

PT Non-radioactive detection of p53 specific antibodies - by capture

PT on immobilised p53 or its fragments, then reaction with labelled

PT second antibody, for diagnosis of tumours and suitable for

PT screening

PS Claim 13; Page 24; 35pp; German.

CC Antibodies specific for p53 are detected by binding to immobilised

CC fragments of the p53 gene product containing the antibody-binding

CC region. Preferred fragments contain amino acids 1-241, 40-349,

CC 40-393, 66-241, 66-393, 237-349, 237-393 and esp. 9-33, 37-52 or

CC 368-386. See Q62357-Q62366 for nucleic acid sequences coding for

CC each of these fragments.

SQ Sequence 984 BP; 224 A; 311 C; 259 G; 190 T;

Query Match 56.7%; Score 17; DB 11; Length 984;

Best Local Similarity 100.0%; Pred. No. 3.91e+00;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 889 agcaggtcactccag 905
|||||

Cp 26 AGCAGGCTCCTCAG 10

RESULT 7
ID Q62359 standard; cDNA; 1062 BP.
AC Q62359;
DT 18-NOV-1994 (first entry)

DE Fragment coding for human p53 amino acids 40-393.

KM Human nuclear phosphoprotein p53; tumour suppressor gene product;

OS anti-oncogene; cancer; tumour; antibody binding region; epitope; ds.

05 Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..1062

FT /tag= a

FT /product= p53(40-393)

FT /transl_except= pos:700..702,aa:Arg

FT /note= "partial CDS, therefore sequence does not

FT contain start or stop codons"

FN M09408241-A.

PD 14-APR-1994.

PF 30-SEP-1993; E02666.

PR 30-SEP-1992; DE-232823.

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM STIFTUNG.

PI Klein R, Schranz P, Tessmer C, Volkmann M, Zentgraf H;

DR WPI; 94-135732/16.

DR P-PSDB; R51874.

PT Non-radioactive detection of p53 specific antibodies - by capture
 PT on immobilised p53 or its fragments, then reaction with labelled
 PT second antibody, for diagnosis of tumours and suitable for
 PT screening
 PS Claim 13; Page 22; 35pp; German.
 CC Antibodies specific for p53 are detected by binding to immobilised
 CC fragments of the p53 gene product containing the antibody-binding
 CC region. Preferred fragments contain amino acids 1-241, 40-349,
 CC 40-393, 66-241, 66-393, 237-349, 237-393 and esp. 9-33, 37-52 or
 CC 368-386. See 062357-062366 for nucleic acid sequences coding for
 CC each of these fragments.
 SQ Sequence 1062 BP; 246 A; 329 C; 279 G; 208 T;

Query Match 56.7%; Score 17; DB 11; Length 1062;
 Best Local Similarity 100.0%; Pred. No. 3.91e+00;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 967 agcagggctcactccag 983
 Cp 26 ACACAGGCTCTACTCCAG 10

RESULT 8

ID Q29632 standard; DNA; 1173 BP.
 AC Q29632;
 DT 16-MAR-1993 (first entry)
 DE Hepatitis C virus HC-J1 3' region.
 KM Non-A non-B hepatitis; NANBH; HCV; detection; diagnosis; screening;
 KM PCR; primer; polymerase chain reaction; ss.
 OS Hepatitis C virus.
 PN EP-510952-A.
 PD 28-OCT-1992.
 PF 23-APR-1992; 303625.
 PR 26-APR-1991; JP-191376.
 PA (IMMO) IMMO NO JAPAN INC.
 PI Nakamura T, Okamoto H;
 DR WPI; 92-359137/44.
 PT Detection of non-A, non-B hepatitis virus - using new
 PT oligo-nucleotide primers with nucleotide sequences corresp. to
 PT part. of the viral RNA
 PS Disclosure; Page 27; 54pp; English.
 CC This sequence represents the 3' region of hepatitis C virus RNA. The
 CC original sample was obtained from human and chimpanzee plasma. RNA
 CC was isolated from several samples and homology compared, and the
 CC respective sequence of about 1900 - 2500 nucleotides of the 5'
 CC terminus and 1100 nucleotides of the 3' terminus determined.
 SQ Sequence 1173 BP; 248 A; 348 C; 305 G; 272 T;

Query Match 56.7%; Score 17; DB 5; Length 1173;
 Best Local Similarity 90.5%; Pred. No. 3.91e+00;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 516 agacagcagcactccag 536
 Cp 30 AGACACAGCGCTCTACTCCAG 10

RESULT 9

ID Q97854 standard; cDNA; 1181 BP.
 AC Q97854;
 DT 06-DEC-1995 (first entry)
 DE Human p53 cDNA.
 KM Ubiquitin-conjugating enzyme; p53 protein; cell cycle;
 KM cell proliferation; cancer; poitiasis; fibrosis; ds.

OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..1181
 FT /*tag= a
 PN W09518974-A.
 PD 13-JUL-1995.
 PF 04-JAN-1995; U00164.
 PR 04-JAN-1994; US-176937.
 PR 23-MAY-1994; US-247904.
 PR 27-MAY-1994; US-250795.
 PR 13-SEP-1994; US-305520.
 PA (MITO-) MITOTIX INC.
 PI Cottarel G, Draetta G, Eckstein JW, Gyuris J, Rolfe M;
 DR WPI; 95-255137/33.
 DR P-P5DB; R79658.
 PT Identifying inhibitors of ubiquitin mediated proteolysis of cell cycle
 PT regulatory proteins - also new ubiquitin conjugating enzymes, their
 PT related nucleic acid, vectors, antibodies etc., useful for regulating
 PT e.g. cell proliferation
 PS Disclosure; Page 105-106; 157pp; English.
 CC Human p53 cDNA (given in Q97854) was amplified from a HeLa cell
 CC cDNA library using the primers given in Q97852-53. The gene
 CC was subcloned into a baculovirus vector for expression of
 CC recombinant p53 in Sf9 insect cells for use as a component of
 CC an in vitro ubiquitin conjugating system.
 SQ Sequence 1181 BP; 275 A; 366 C; 306 G; 234 T;

Query Match 56.7%; Score 17; DB 15; Length 1181;
 Best Local Similarity 100.0%; Pred. No. 3.91e+00;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1084 agcagggctcactccag 1100
 Cp 26 ACACAGGCTCTACTCCAG 10

RESULT 10

ID T29719 standard; cDNA; 1182 BP.
 AC T29719;
 DT 29-OCT-1996 (first entry)
 DE Wild type p53 gene sequence.
 KM p53 gene; cancer; carcinoma; neoplastic; neoplasia; phenotype;
 KM osteosarcoma cells; lung carcinoma cells; lymphoma cells;
 KM leukaemia cells; soft tissue sarcoma cells; breast cells;
 KM bladder cells; prostate carcinoma cell; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..1182
 FT /*tag= a
 FT /product= p53 protein.
 FT misc difference 19..21
 FT /*tag= b
 FT /transl_except= CAT encodes Aspartic acid.
 PN EP-710722-A1.
 PD 08-MAY-1996.
 PF 23-AUG-1991; 307791.
 PR 24-AUG-1990; US-573405.
 PA (REGC) UNIV CALIFORNIA.
 PI Chen P, Lee W;
 DR WPI; 96-223439/23.
 DR P-P5DB; R91933.
 PT Use of wild-type p53 gene - in a medicament for suppressing the
 PT neoplastic phenotype of a cancer cell lacking wild-type p53 protein
 PT Claim 1; Page 5; 25pp; English.

CC The wild-type p53 gene can be used in the production of a medicament
CC for suppressing the neoplastic phenotype of a cancer cell lacking
CC endogenous wild type p53 protein. Cancer cells suppressed in such
CC fashion include osteosarcoma cells, lung carcinoma cells, lymphoma
CC cells, leukemia cells, soft tissue sarcoma cells or breast, bladder
CC or prostate carcinoma cells.
SQ Sequence 1182 BP; 278 A; 366 C; 304 G; 234 T;

Query Match 56.7%; Score 17; DB 22; Length 1182;
Best Local Similarity 100.0%; Pred. No. 3.91e+00;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1084 agcaggctcactccag 1100
|||||
Cp 26 AGCAGGCTCCTCCTCAG 10

RESULT 11
ID 067884 standard; DNA; 1182 BP.
AC 067884;

DT 23-MAR-1995 (first entry)
DE Human p53 DNA.
KM Polymerase chain reaction; primer; amplify; NYVAC; ALVAC; recombinant;
KM murine; interleukin-2; IL-2; pRM825; pmut-1; pBS-SK; pM151; TK vector;
KM plasmid; vaccinia; H6 promoter; amplify; primer; antigenic response;
KM polymerase chain reaction; poxvirus; pSD542; immunological response;
KM pathogen; human; interferon; IFN, ss.
OS Synthetic.
PN W09416716-A.

PD 04-AUG-1994.
PF 21-JAN-1994; 000888.
PR 21-JAN-1993; US-007115.
PR 19-JAN-1994; US-184009.
PA (VIR0-) VIRGENETICS CORP.
PI Cox WI, Paolletti E, Tartaglia J;
DR WPI; 94-263767/32.

PT DNA encoding cytokine and/or tumour associated antigen
PS Example 32; Fig 39; 232pp; English.

CC This sequence represents the wildtype human p53 gene from the translation
CC initiation codon to the stop codon. This sequence was used in the
CC construction of an ALVAC-based recombinant virus containing a mutant
CC form of the human p53 gene. The mutant form has a G>A substitution at
CC position 524, changing an Arg residue at position 175 to a His residue.
CC The plasmid pM110 (see also 067864) contains the vaccinia H6 promoter
CC and the wild type human p53 gene in the ALVAC C5 insertion site. The
CC mutant p53 gene was obtained from plasmid Cx22A and cloned into pM110
CC to generate pM143. Recombination between pM143 and ALVAC rescuing
CC virus produced recombinant virus vCP270, which contains the vaccinia H6
CC promoted mutated human p53 in the C5 locus. The resulting virus may be
CC used in a composition for inducing an antigenic or immunological
CC response, ie. for immunisation against pathogene.

SQ Sequence 1182 BP; 276 A; 365 C; 307 G; 234 T;
Query Match 56.7%; Score 17; DB 12; Length 1182;
Best Local Similarity 100.0%; Pred. No. 3.91e+00;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1084 agcaggctcactccag 1100
|||||
Cp 26 AGCAGGCTCCTCCTCAG 10

RESULT 12

ID 022995 standard; DNA; 1182 BP.
AC 022995;
DT 23-JUL-1992 (first entry)
DE Sequence encoding 53 kD cellular protein.
KM Cancer therapy; cancer suppressor gene; oncogenesis; ss.
OS Homo sapiens.

FH Key Location/Qualifiers
FT CDS 1..1182

FT /tag= a
PN EP-475623-A.
PD 18-MAR-1992.
PF 23-AUG-1991; 307791.
PR 24-AUG-1990; US-573405.
PA (RECC) UNIV OF CALIFORNIA.

P1 Lee WH, Chen PL;
DR WPI; 92-090221/12.
DR P-PSDB; R22238.
PT Cloned p53 cDNA and protein prods. - for suppression of
PT neoplastic phenotype e.g. in osteo-sarcoma(s), leukaemia(s),
PT lymphoma(s), etc.
PS Claim 5; Page 15; 25pp; English.

CC p53 cDNA, or its gene prods., can be used to suppress and eradicate
CC cancers caused by defective, mutant or absent cancer suppressor
CC genes. Variant forms of p53 are found in human breast, lung or
CC colon carcinoma, lymphoma, leukaemia, etc., suggesting that mutation
CC of the p53 genes is involved in oncogenesis. Specifically 273 Arg
CC is replaced by 273 His, a mutation found exclusively in tumour cells.
SQ Sequence 1182 BP; 277 A; 368 C; 303 G; 234 T;

Query Match 56.7%; Score 17; DB 3; Length 1182;
Best Local Similarity 100.0%; Pred. No. 3.91e+00;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1084 agcaggctcactccag 1100
|||||
Cp 26 AGCAGGCTCCTCCTCAG 10

RESULT 13

ID 067864 standard; DNA; 1483 BP.
AC 067864;

DT 22-MAR-1995 (first entry)
DE Vaccinia H6 promoter/human p53 expression cassette from vCP207.
KM Polymerase chain reaction; primer; amplify; NYVAC; ALVAC; recombinant;
KM human; p53; wild type; mutant; p53wtcbaisp6/T3; p53-21Xba1; p53-238Xba1;
KM vaccinia; H6 promoter; plasmid; pRM825; expression cassette;
KM ALVAC(CPpp) p53; poxvirus; antigenic response; immunological response;
KM pathogen; ss.
OS Synthetic.

FH Key Location/Qualifiers
FT misc_feature 1..108
FT /tag= a
FT /note= "Flanking sequence"
FT promoter 109..232
FT /tag= b
FT /note= "Vaccinia H6 promoter"
FT CDS 233..1414

FT /tag= c
FT /product= Wild type human p53
FT misc_feature 1415..1483
FT /tag= d
FT /note= "Flanking sequence"
PN W09416716-A.
PD 04-AUG-1994.

PF 21-JAN-1994; U00888.
PR 21-JAN-1993; US-007115.
PR 19-JAN-1994; US-184009.
PA (VIR0-) VIRGENETICS CORP.
PI Cox MI, Paoletti E, Tartaglia J;
PI WPI; 94-263767/32.
PT Attenuated recombinant virus used for cancer therapy - comprises
PT DNA encoding cytokine and/or tumour associated antigen
PS Example 15; Fig 18; 232pp; English.
CC The sequences given in 067863-64 represent expression cassettes
CC containing the vaccinia H6 promoter and the wild type human p53 gene
CC in VP1101 and vCP207, respectively. These sequences were used in the
CC construction of poxvirus based recombinant viruses expressing wild type
CC and mutant forms of the human p53 gene product. The wild type and
CC mutant genes were derived from plasmids p53wtXbaISp6/T3, p53-21XbaI and
CC p53-238XbaI respectively. The vaccinia H6 promoter was derived from
CC plasmid pRM825 and precisely linked to the 5'-most region of the p53
CC gene. The expression cassettes generated by linking the vaccinia H6
CC promoter and the p53 genes may be cloned in to the poxvirus insertion
CC plasmid, AlVAC, to give vCP207, vCP193 and vCP191 recombinant
CC viruses, and into NYVAC to give VP1101, VP1096 and VP1098. These
CC viruses may be used in a composition for inducing an antigenic or
CC immunological response, ie. for immunisation against pathogens.
SQ Sequence 1483 BP; 367 A; 415 C; 372 G; 329 T;

Query Match 56.7%; Score 17; DB 12; Length 1483;
Best Local Similarity 100.0%; Pred. No. 3.91e+00;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1316 agcagggctcactccag 1332
|||||
Cp 26 ACACAGGCTCCTCCTCAG 10

RESULT 14
ID 067863 standard; DNA; 1512 BP.
AC 067863;
DT 22-MAR-1995 (first entry)
DE Vaccinia H6 promoter/human p53 expression cassette from VP1101.
KW Polymerase chain reaction; primer; amplify; NYVAC; recombinant;
KW human; p53; wild type; mutant; p53wtXbaISp6/T3; p53-21XbaI; p53-238XbaI;
KW vaccinia; H6 promoter; plasmid; pRM825; expression cassette;
KW AlVAC(CPP)p53; poxvirus; antigenic response; immunological response;
KW pathogen; ss.
OS Synthetic.
FH Key
FT misc_feature 1..144
FT /tag= a
FT /note= "Flanking sequence"
FT promoter 145..268
FT /tag= b
FT /note= "Vaccinia H6 promoter"
FT CDS 269..1450
FT /tag= c
FT /product= Human wildtype p53
FT misc_feature 1451
FT /tag= d
FT /note= "Flanking sequence"
PN M09416716-A.
PD 04-AUG-1994.
PF 21-JAN-1994; U00888.
PR 21-JAN-1993; US-007115.
PR 19-JAN-1994; US-184009.
PA (VIR0-) VIRGENETICS CORP.

PI Cox MI, Paoletti E, Tartaglia J;
PI WPI; 94-263767/32;
PT Attenuated recombinant virus used for cancer therapy - comprises
PT DNA encoding cytokine and/or tumour associated antigen
PS Example 15; Fig 17; 232pp; English.
CC The sequences given in 067863-64 represent expression cassettes
CC containing the vaccinia H6 promoter and the wild type human p53 gene
CC in VP1101 and vCP207, respectively. These sequences were used in the
CC construction of poxvirus based recombinant viruses expressing wild type
CC and mutant forms of the human p53 gene product. The wild type and
CC mutant genes were derived from plasmids p53wtXbaISp6/T3, p53-21XbaI and
CC p53-238XbaI respectively. The vaccinia H6 promoter was derived from
CC plasmid pRM825 and precisely linked to the 5'-most region of the p53
CC gene. The expression cassettes generated by linking the vaccinia H6
CC promoter and the p53 genes may be cloned in to the poxvirus insertion
CC plasmid, AlVAC, to give vCP207, vCP193 and vCP191 recombinant
CC viruses, and into NYVAC to give VP1101, VP1096 and VP1098. These
CC viruses may be used in a composition for inducing an antigenic or
CC immunological response, ie. for immunisation against pathogens.
SQ Sequence 1512 BP; 379 A; 419 C; 380 G; 334 T;

Query Match 56.7%; Score 17; DB 12; Length 1512;
Best Local Similarity 100.0%; Pred. No. 3.91e+00;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1352 agcagggctcactccag 1368
|||||
Cp 26 ACACAGGCTCCTCCTCAG 10

RESULT 15
ID 074770 standard; cDNA; 9502 BP.
AC 074770;
DT 01-AUG-1995 (first entry)
DE Hepatitis C virus gene HC-J1/cDNA.
KW Hepatitis C virus; HCV gene HC-J1/cDNA; specific antibodies; ss.
OS Hepatitis C virus.
FH Key
FT CDS 342..9377
FT /tag= a
FT EN 306284887-A.
PD 11-OCT-1994.
PF 10-DEC-1993; 345753.
PR 10-DEC-1992; JP-360705.
PR (IMMO) IMMONO JAPAN KK.
PA WPI; 94-362594/45.
DR P-PSDB; R66995.
DE HCV genes and the corresponding proteins - used in the production
PT of anti-HCV antibodies and the detection of HCV infection
PS Claim 7; Pages 12-17; 35pp; Japanese.
CC 074770 encodes R66995 the HC-J1/protein, the cDNA can be used in
CC the construction of an expression vector for the transformation
CC of a host cell. The host cell can then be used in the production
CC of proteins and peptides, useful in the preparation of monoclonal
CC and polyclonal HCV-specific antibodies.
SQ Sequence 9502 BP; 1879 A; 2847 C; 2691 G; 2085 T;

Query Match 56.7%; Score 17; DB 13; Length 9502;
Best Local Similarity 90.5%; Pred. No. 3.91e+00;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 8794 agacagcagcactccag 8814
|||||
Cp 30 ACACAGGCTCCTCCTCAG 10

Mar 25 02:50

US-08-644-289-4.rmg

13

Search completed: Tue Mar 25 02:52:51 1997
Job time : 15 secs.

US-08-644-289-4 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Match on: n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Mar 25 02:51:15 1997; Maspar time 56.88 Seconds

Tabular output not generated.

Title: >US-08-644-289-4
Description: (1-30) from US08644289.seq
Perfect Score: 30
N.A. Sequence: 1 AGTCGATCCTCGAGTACGACCTGCTCTCT 30
Comp: TCAGCCTAGACCTCCTCGAGCAGACAGA

Scoring table: TABLE default
Gap 10

Mismatch STD: Dbase 0; Query 0

Searched: 279077 seqs, 411808665 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb1-new5
1:BCT 2:FTON 3:INV1 4:INV2 5:ORC 6:MAM 7:VRT 8:PLN 9:PRI

Database: genbank94
10:PRO1 11:PRO2 12:ROD 13:SYN 14:UNC 15:VIR

16:BCT1 17:BCT2 18:BCT3 19:BCT4 20:BCF5 21:BCT6 22:BCT7
23:BCT8 24:BCT9 25:INV1 26:INV2 27:INV3 28:INV4 29:INVS
30:INV6 31:INV7 32:MAM1 33:MAM2 34:MAM3 35:VRT1 36:VRT2
37:VRT3 38:PAT1 39:PAT2 40:PAT3 41:PHG 42:PLN1 43:PLN2
44:PLN3 45:PLN4 46:PLN5 47:PLN6 48:PLN7 49:PLN8 50:PRI1
51:PRI2 52:PRI3 53:PRI4 54:PRI5 55:PRI6 56:PRI7 57:PRI8
58:PRI9 59:PRI10 60:PRI11 61:PRI12 62:PRI13 63:ROD1
64:ROD2 65:ROD3 66:ROD4 67:ROD5 68:ROD6 69:ROD7 70:ROD8
71:STR 72:SYN 73:UNA 74:VRL1 75:VRL2 76:VRL3 77:VRL4
78:VRL5 79:VRL6 80:VRL7 81:VRL8

Database: genbank-new5
82:BCT 83:INV1 84:INV2 85:MAM 86:VRT 87:PAT 88:PHG
89:PLN 90:PRI 91:ROD 92:STR 93:SYN 94:UNA 95:VRL
Database: u-emb146 94
96:part1

Statistics: Mean 7.318; Variance 2.910; scale 2.515

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	22	73.3	159 69	RATP53TS07	Rattus norvegicus tum	1.43e+03
2	22	73.3	1322 67	MUSP53C	Mouse p53 mRNA, compl	1.43e+03
3	21	70.0	107 63	MMAN10	Murine gene for cellu	8.05e+03
4	21	70.0	1241 67	MUSP53B	Mouse p53 mRNA, compl	8.05e+03
5	21	70.0	1285 67	MUSP53A	Mouse p53 mRNA, compl	8.05e+03
6	21	70.0	1377 64	MMP53	Mouse mRNA for trans	8.05e+03
7	21	70.0	1447 91	RATP53C	Rattus norvegicus p53	8.05e+03
8	21	70.0	1447 69	RATP53D	Rattus norvegicus p53	8.05e+03
9	21	70.0	1447 12	RNP53E	Rattus norvegicus p53	8.05e+03
10	21	70.0	1621 69	RNP53	Rat mRNA for nuclear	8.05e+03
11	21	70.0	1772 67	MUSP53M	Mouse p53 cellular tu	8.05e+03
12	21	70.0	1773 64	MMP53R	Mouse mRNA for cellu	8.05e+03
13	21	70.0	2130 67	MUSP53G	Mouse p53 cellular tu	8.05e+03
14	21	70.0	2132 64	MMP53P	Mouse pseudogene for	8.05e+03
15	20	66.7	1273 70	RNU07019	Rattus norvegicus Mts	4.36e+02
16	18	60.0	2633 69	RATP53REC	Human phosphoprotein	1.13e+00
17	18	60.0	20303 54	HSP53G	Human p53 gene for tr	1.13e+00
18	18	60.0	54336 61	HOMAPDGL	Homo sapiens ERCC2 ge	1.13e+00
19	18	60.0	54336 9	HSP53G	Homo sapiens ERCC2 ge	1.13e+00
20	18	60.0	844 33	DOG53MUT	Canis familiaris p53	5.32e+00
21	17	56.7	1179 54	HSP53010	Human mRNA for mutate	5.32e+00
22	17	56.7	1179 54	HSP53006	Human mRNA for mutate	5.32e+00
23	17	56.7	1179 54	HSP53005	Human mRNA for mutate	5.32e+00
24	17	56.7	1179 54	HSP53011	Human mRNA for mutate	5.32e+00
25	17	56.7	1179 54	HSP53004	Human mRNA for mutate	5.32e+00
26	17	56.7	1179 54	HSP53003	Human mRNA for mutate	5.32e+00
27	17	56.7	1179 54	HSP53002	Human mRNA for mutate	5.32e+00
28	17	56.7	1179 54	HSP53001	Human mRNA for mutate	5.32e+00
29	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
30	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
31	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
32	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
33	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
34	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
35	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
36	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
37	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
38	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
39	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
40	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
41	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
42	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
43	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
44	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00
45	17	56.7	1179 54	HSP53000	Human mRNA for mutate	5.32e+00

ALIGNMENTS

RESULT	1	RATP53TS07	159 bp	DNA	ROD	14-JUL-1993
LOCUS						
DEFINITION		Rattus norvegicus tumor suppressor (p53) gene, exon 9.				
ACCESSION		L07909 L07781				
NID		g205949				
KEYWORDS		tumor suppressor.				
SEGMENT		7 of 8				
SOURCE		Rattus norvegicus (strain Sprague-Dawley) DNA.				
ORGANISM		Rattus norvegicus				

Mar 25 02:50

US-08-644-289-4.rge

3

Eukaryotes; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.

REFERENCE 1 (bases 1 to 159)
AUTHORS Hulla, J.E. and Schneider, R.P.St..
TITLE Structure of the rat p53 tumor suppressor gene
JOURNAL Nucleic Acids Res. 21, 713-717 (1993)
MEDLINE 93181268
FEATURES
source location/Qualifiers
1..159
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/sequenced_mol="DNA"
order(107908:315..340,1..26)
/gene="p53"
/note="Intron 8 is ~830 bp"
/number=8
27..133
/gene="p53"
/number=9
BASE COUNT 32 a 40 c 53 g 34 t
ORIGIN

Query Match 73.3%; Score 22; DB 69; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.43e-03;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 113 agacagcaggctcactccagc 134
|||||
Cp 30 AGACAGCAGCGCTCCTCAGC 9

RESULT 2
LOCUS M13874 1322 bp mRNA ROD 02-NOV-1992
DEFINITION Mouse p53 mRNA, complete cds, clone p53-m8.
ACCESSION M13874
KEYWORDS p53 gene.
SOURCE Mouse Ab-MuLV transformed Meth A fibroblast and normal T-cell line L12, cDNA to mRNA, clone p53-m8.
ORGANISM Mus musculus
Eukaryotes; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE 1 (bases 1 to 1322)
AUTHORS Arai, N., Nomura, D., Yokota, K., Wolf, D., Brill, E., Shohat, O. and Rotter, V.
TITLE Immunologically distinct p53 molecules generated by alternative splicing
JOURNAL Mol. Cell. Biol. 6, 3232-3239 (1986)
MEDLINE 87064640
REFERENCE 2 (bases 1 to 1322)
AUTHORS Han, K.A. and Kulesz-Martin, M.F.
TITLE Alternatively spliced p53 mRNA in transformed and normal cells of different tissue types
JOURNAL Nucleic Acids Res. 20, 1979-1981 (1992)
MEDLINE 92253421
FEATURES
source location/Qualifiers
1..1322
/organism="Mus musculus"
/cell_line="normal T-cell line L12"
/cell_type="Ab-MuLV transformed Meth A fibroblast"
/sequenced_mol="cDNA to mRNA"
54..119
/gene="p53"
/map="11"

CDS

Mar 25 02:50

US-08-644-289-4.rge

4

/codon_start=1
/db_xref="PID:g200203"
/translation="MTAMEESQDISLELPDSQFTSGMLKLPDPDILPSPHOMDL
LLEQDVEFEFGSEALAVSGAPRAADPVETGVPAPAPATPRPLSSVPSQKTTQG
NYGFHLGIAQGTAKSVACTYSPPLKLEFQLAKTCPOVLWVSTPPAGSRVRAAMLY
KKSQMTVVRQCPHHERGSDGGLAPQHLIRVENIAPPEYLEDQTRISVVPYE
PPAGSEYTTIHYKVMKSSCGMGMRRLITITILEDSNLLGDSFEVRVACPG
RDRRTEDENRKEKVLCEPLPGSARALPTCTASPPQKKKPLDGEYFTLKIGRRR
FEMREINFALEIKDHAHTEESGDSHSSIDPPARQALIKESPPNC"

BASE COUNT 308 a 407 c 329 g 278 t
ORIGIN

Query Match 73.3%; Score 22; DB 67; Length 1322;
Best Local Similarity 92.3%; Pred. No. 1.43e-03;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1124 agacagcaggctcactccagctcc 1149
|||||
Cp 30 AGACAGCAGCGCTCCTCAGCATCC 5

RESULT 3
LOCUS M13874 107 bp DNA ROD 16-JUN-1995
DEFINITION Murine gene for cellular tumour antigen p53 (exon 10).
ACCESSION X00884
KEYWORDS antigen.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 107)
AUTHORS Bienz, R., Zakut-Houri, R., Givol, D. and Oren, M.
TITLE Analysis of the gene coding for the murine cellular tumour antigen p53
JOURNAL EMBO J. 3 (9), 2179-2183 (1984)
MEDLINE 85027173
REFERENCE 2 (bases 1 to 107)
AUTHORS Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Givol, D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306 (5943), 594-597 (1983)
MEDLINE 84068204
FEATURES
source location/Qualifiers
1..107
/organism="Mus musculus"
exon 10
/label=ex10
/usedin=X00876:p53_CDS
/usedin=X00876:p53_mRNA
BASE COUNT 27 a 25 c 35 g 20 t
ORIGIN

Query Match 70.0%; Score 21; DB 63; Length 107;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 87 agacagcaggctcactccagc 107
|||||
Cp 30 AGACAGCAGCGCTCCTCAGC 10

RESULT 4
LOCUS M13874 1241 bp mRNA ROD 02-NOV-1992

DEFINITION Mouse p53 mRNA, complete cds, clone p53-m1.
ACCESSION M13873
NID 9200200
KEYWORDS p53 gene.
SOURCE Mouse (BALB/c) Meth A library, cDNA to mRNA, clone p53-m1.
ORGANISM Mus musculus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE
AUTHORS Arai,N., Nomura,D., Yokota,K., Wolf,D., Brill,E., Shohat,O. and Rotter,V.
TITLE Immunologically distinct p53 molecules generated by alternative splicing
JOURNAL Mol. Cell. Biol. 6, 3232-3239 (1986)
MEDLINE 87064640
FEATURES
source Location/Qualifiers
1..1241
/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/sequenced_mol="cDNA to mRNA"
/issue_1lb="Meth A"
69..1241
/gene="p53"
/map="11"
/codon_start=1
/db_xref="PID:g200201"
/translation="MTAMEESQSDISLELPISQETSGIMKLPPEDILPSHCMDL
LLPDVVEEFEGSEALRVSGAPADDPVTEPGVPAPATPMPILSSFPVSOQTYOG
NYGFHLGFIQSGTAKSYMCTYSPPLKLFQCLAKTCYQULWVSAAPPAGSRVRAAIY
KKSQHMTVVARCPHERCSDDGLAPPHILINVEGNLPEYLEDROTFRHSVVPYVE
PEAGSEYTTIHYKYMNSSCMGGMNRPIILTIITLEDSSGNLGRDSEVAVACPG
RDRTEENFRKEVLCPELPFGSAKRALPTCTSAAPQKKPDLDEGYTLKIRGRF
FEMRELNEALELKDAAITEESGDSRAHSSYLTKKGQSTSRHKKTWKKVGPDSD"

BASE COUNT 289 a 376 c 321 g 255 t
ORIGIN

Query Match 70.0%; Score 21; DB 67; Length 1241;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1139 agacagcaggctcactccag 1159
|||||
Cp 30 AGACAGCAGGCTCCTCCTCAG 10

RESULT 5
LOCUS MUSP53A 1285 bp mRNA ROD 05-NOV-1992
DEFINITION Mouse p53 mRNA, complete cds, clone pcd53.
ACCESSION M13872
NID 9200198
KEYWORDS p53 gene.
SOURCE Mouse (BALB/c) nontransformed helper T-cell, cDNA to mRNA, clone pCD-p53.
ORGANISM Mus musculus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE
AUTHORS Arai,N., Nomura,D., Yokota,K., Wolf,D., Brill,E., Shohat,O. and Rotter,V.
TITLE Immunologically distinct p53 molecules generated by alternative splicing
JOURNAL Mol. Cell. Biol. 6, 3232-3239 (1986)
MEDLINE 87064640

FEATURES
source Location/Qualifiers
1..1285
/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/cell_type="nontransformed helper T-cell"
/sequenced_mol="cDNA to mRNA"
113..1285
/codon_start=1
/product="p53 protein"
/db_xref="PID:g200199"
/translation="MTAMEESQSDISLELPISQETSGIMKLPPEDILPSHCMDL
LLPDVVEEFEGSEALRVSGAPADDPVTEPGVPAPATPMPILSSFPVSOQTYOG
NYGFHLGFIQSGTAKSYMCTYSPPLKLFQCLAKTCYQULWVSAAPPAGSRVRAAIY
KKSQHMTVVARCPHERCSDDGLAPPHILINVEGNLPEYLEDROTFRHSVVPYVE
PEAGSEYTTIHYKYMNSSCMGGMNRPIILTIITLEDSSGNLGRDSEVAVACPG
RDRTEENFRKEVLCPELPFGSAKRALPTCTSAAPQKKPDLDEGYTLKIRGRF
FEMRELNEALELKDAAITEESGDSRAHSSYLTKKGQSTSRHKKTWKKVGPDSD"

BASE COUNT 299 a 385 c 334 g 267 t
ORIGIN

Query Match 70.0%; Score 21; DB 67; Length 1285;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1183 agacagcaggctcactccag 1203
|||||
Cp 30 AGACAGCAGGCTCCTCCTCAG 10

RESULT 6
LOCUS MMP53 1377 bp RNA ROD 12-SEP-1993
DEFINITION Mouse mRNA for transformation associated protein p53.
ACCESSION X00741
NID 93570
KEYWORDS oncogene.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires; Rodentia;
Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
1 (bases 1 to 1377)
Jenkins,J.R., Rudge,K., Redmond,S. and Wade-Evans,A.
Cloning and expression analysis of full length mouse cDNA sequences
encoding the transformation associated protein p53
Nucleic Acids Res. 12 (14), 5609-5626 (1984)
JOURNAL 84272240
MEDLINE
COMMENT Data kindly reviewed (19-FEB-1986) by A. Wade-Evans.
FEATURES
source Location/Qualifiers
1..1377
/organism="Mus musculus"
123..1292
/note="p53"
/codon_start=1
/db_xref="PID:g53571"
/translation="MTAMEESQSDISLELPISQETSGIMKLPPEDILPSHCMDL
LLPDVVEEFEGSEALRVSGAPADDPVTEPGVPAPATPMPILSSFPVSOQTYOG
NYGFHLGFIQSGTAKSYMCTYSPPLKLFQCLAKTCYQULWVSAAPPAGSRVRAAIY
KKSQHMTVVARCPHERCSDDGLAPPHILINVEGNLPEYLEDROTFRHSVVPYVE
PEAGSEYTTIHYKYMNSSCMGGMNRPIILTIITLEDSSGNLGRDSEVAVACPG
RDRTEENFRKEVLCPELPFGSAKRALPTCTSAAPQKKPDLDEGYTLKIRGRF
FEMRELNEALELKDAAITEESGDSRAHSSYLTKKGQSTSRHKKTWKKVGPDSD"

Mar 25 02:50

US-08-644-289-4.rge

7

BASE COUNT 316 a 422 c 348 g 291 t
ORIGIN
Query Match 70.0%; Score 21; DB 64; Length 1377;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1190 agacagcaggtcactccag 1210
|||||
Cp 30 AGACAGCAGCGCTCCTCCAG 10
RESULT 7
LOCUS R147PSEUDO 1447 bp DNA ROD 12-APR-1996
DEFINITION Rattus norvegicus p53 (PG-III) pseudogene, partial ORF.
ACCESSION L12046
NID g206472
KEYWORDS tumor suppressor.
SOURCE Rattus norvegicus (strain Fischer 344) DNA.
ORGANISM Rattus norvegicus
Eukaryota; Eutheria; Rodentia; Muridae;
Murinae; Rattus.
REFERENCE 1 (bases 1 to 1447)
AUTHORS Meghori, C.M., Buzard, G.S., Calvert, R.J., Hulla, J.E. and Rice, J.M.
TITLE Cloning and sequence of a processed p53 pseudogene from rat: a
JOURNAL potential source of false 'mutations' in PCR fragments of tumor DNA
MEDLINE Gene 166 (2), 317-322 (1995)
COMMENT 96125211
FEATURES
source location/Qualifiers
1..1447
/organism="Rattus norvegicus"
/strain="Fischer 344"
/sex="male"
/tissue_type="liver"
1..75
/note="this region of the pseudogene corresponds to exon 2
of the rat p53 cDNA; putative"
misc_signal 2..4
/note="this site within the pseudogene corresponds to the
ATC translation start site of the rat p53 cDNA; putative"
76..97
/note="this region of the pseudogene corresponds to exon 3
of the rat p53 cDNA; putative"
98..315
/note="this region of the pseudogene corresponds to exon 4
of the rat p53 cDNA; putative"
CDS <168..>542
/gene="p53 PG-III"
/note="tumor suppressor; this region of the pseudogene is
a potential open reading frame; putative; NCBI gi:
206473"
/codon_start=1
/db_xref="PID:g206473"
/translation="KVGRKSKQQLP HNNLELPISLTKNLSQLTSGLPVSDNQ
VCYHVPSPKRIATLPAGEDMPCAVMQQLHTSNMHLACAGCIYKSGHMEVARCH
HERCSDDDDQTPPPPTPSISG"
316..498
/gene="p53 PG-III"
/note="this region of the pseudogene corresponds to exon 5
of the rat p53 cDNA; putative"
499..746
/note="this region of the pseudogene corresponds to exon 6
exon
exon

Mar 25 02:50

US-08-644-289-4.rge

8

exon of the rat p53 cDNA; putative"
747..882
/note="this region of the pseudogene corresponds to exon 7
of the rat p53 cDNA; putative"
883..937
exon /note="this region of the pseudogene corresponds to exon 8
of the rat p53 cDNA; putative"
938..1031
/note="this region of the pseudogene corresponds to exon 9
of the rat p53 cDNA; putative"
1032..1447
exon /note="this region of the pseudogene corresponds to exon
10 of the rat p53 cDNA; putative"
1107..1109
misc_signal /note="this site within the pseudogene corresponds to the
translation stop site of the rat p53 cDNA; putative"
BASE COUNT 366 a 388 c 336 g 357 t
ORIGIN
Query Match 70.0%; Score 21; DB 91; Length 1447;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1011 agacagcaggtcactccag 1031
|||||
Cp 30 AGACAGCAGCGCTCCTCCAG 10
RESULT 8
LOCUS R147PSEUDO 1447 bp DNA ROD 07-JUN-1993
DEFINITION Rattus norvegicus p53 tumor suppressor pseudogene, partial ORF.
ACCESSION L12046
NID g206472
KEYWORDS tumor suppressor.
SOURCE Rattus norvegicus (strain Fisher 344) male liver DNA.
ORGANISM Rattus norvegicus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Muridae; Murinae.
REFERENCE 1 (bases 1 to 1447)
AUTHORS Meghori, C.M., Buzard, G.S., Hulla, J.E., Calvert, R.J. and Rice, J.M.
TITLE Nucleotide sequence of a rat p53 pseudogene: a representative of
JOURNAL the family of p53 pseudogenes in the rat genome
Unpublished (1993)
FEATURES
source location/Qualifiers
1..1447
/organism="Rattus norvegicus"
/strain="Fisher 344"
/sequenced_mol="DNA"
/sex="male"
/tissue_type="liver"
1..75
/note="this region of the pseudogene corresponds to exon 2
of the rat p53 cDNA; putative"
2..4
/note="this site within the pseudogene corresponds to the
ATC translation start site of the rat p53 cDNA; putative"
76..97
exon /note="this region of the pseudogene corresponds to exon 3
of the rat p53 cDNA; putative"
98..315
/note="this region of the pseudogene corresponds to exon 4
of the rat p53 cDNA; putative"
CDS <168..>542
/note="this region of the pseudogene is a potential open
misc_signal 2..4
/note="this site within the pseudogene corresponds to the
ATC translation start site of the rat p53 cDNA; putative"
76..97
exon /note="this region of the pseudogene corresponds to exon 3
of the rat p53 cDNA; putative"
98..315
/note="this region of the pseudogene corresponds to exon 4
of the rat p53 cDNA; putative"
CDS <168..>542
/note="this region of the pseudogene is a potential open
exon
exon

```
reading frame; putative"
/codon_start=1
/db_xref="PID:g206473"
/translation="KVGRKPSKQQLPHRNLELRPLSLKKNLSQMLSSGLPANSNQ
VCYVHLPSPKLIATIPAGDNPCAVMGQQLTSMHLCACHGCIYKKSQHMTVEVRCSSH
HERCSDGDDQTPPPPTPSILSG"
316..498
/note="this region of the pseudogene corresponds to exon 5
of the rat p53 cDNA; putative"
exon
499..746
/note="this region of the pseudogene corresponds to exon 6
of the rat p53 cDNA; putative"
747..882
/note="this region of the pseudogene corresponds to exon 7
of the rat p53 cDNA; putative"
exon
883..937
/note="this region of the pseudogene corresponds to exon 8
of the rat p53 cDNA; putative"
938..1031
/note="this region of the pseudogene corresponds to exon 9
of the rat p53 cDNA; putative"
exon
1032..1447
/note="this region of the pseudogene corresponds to exon
10 of the rat p53 cDNA; putative"
misc_signal
1107..1109
/note="this site within the pseudogene corresponds to the
translation stop site of the rat p53 cDNA; putative"
BASE COUNT      366 a      388 c      336 g      357 t
ORIGIN

Query Match      70.0%; Score 21; DB 69; Length 1447;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1011 agacagcagggtcactccag 1031
|||||
Cp 30 AGACAGCAGCGCTCCTCCAG 10

RESULT 9
ID RNP5UEDO standard; DNA; ROD; 1447 BP.
AC L12046;
DT 16-MAR-1993 (Rel. 35, Created)
DT 14-APR-1996 (Rel. 47, Last updated, Version 3)
DE Rattus norvegicus p53 (PG-III) pseudogene, partial ORF.
KM tumor suppressor.
OS Rattus norvegicus (rat)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
RN [1]
RP 1-1447
RA Meghoret C.M., Buzard G.S., Calvert R.J., Hulla J.E., Rice J.M.;
RT "Cloning and sequence of a processed p53 pseudogene from rat: a
RT potential source of false 'mutations' in PCR fragments of tumor
RT DNA";
RL Gene 166:317-322 (1995).
CC NCBI gi: 206472
FH Key Location/Qualifiers
FT source 1..1447
FT /organism="Rattus norvegicus"
FT /strain="Fischer 344"
FT /sex="male"
FT /tissue_type="liver"
```

```
FT exon 1..75
FT /note="this region of the pseudogene corresponds to exon 2
FT of the rat p53 cDNA; putative"
FT misc_signal 2..4
FT /note="this site within the pseudogene corresponds to the
FT ATG translation start site of the rat p53 cDNA; putative"
FT exon 76..97
FT /note="this region of the pseudogene corresponds to exon 3
FT of the rat p53 cDNA; putative"
FT exon 98..315
FT /note="this region of the pseudogene corresponds to exon 4
FT of the rat p53 cDNA; putative"
FT CDS <168..>542
FT /gene="p53 PG-III"
FT /note="tumor suppressor; this region of the pseudogene is
a potential open reading frame; putative; NCBI gi: 206473"
FT /codon_start=1
FT /db_xref="PID:g206473"
FT /translation="KVGRKPSKQQLPHRNLELRPLSLKKNLSQMLSSGLPANSNQ
V
CYVHLPSPKLIATIPAGDNPCAVMGQQLTSMHLCACHGCIYKKSQHMTVEVRCSSH
E
RCSDGDDQTPPPPTPSILSG"
FT exon 316..498
FT /gene="p53 PG-III"
FT /note="this region of the pseudogene corresponds to exon 5
FT of the rat p53 cDNA; putative"
FT exon 499..746
FT /note="this region of the pseudogene corresponds to exon 6
FT of the rat p53 cDNA; putative"
FT exon 747..882
FT /note="this region of the pseudogene corresponds to exon 7
FT of the rat p53 cDNA; putative"
FT exon 883..937
FT /note="this region of the pseudogene corresponds to exon 8
FT of the rat p53 cDNA; putative"
FT exon 938..1031
FT /note="this region of the pseudogene corresponds to exon 9
FT of the rat p53 cDNA; putative"
FT exon 1032..1447
FT /note="this region of the pseudogene corresponds to exon 10
FT of the rat p53 cDNA; putative"
FT misc_signal 1107..1109
FT /note="this site within the pseudogene corresponds to the
FT translation stop site of the rat p53 cDNA; putative"
SQ Sequence 1447 BP; 366 A; 388 C; 336 G; 357 T; 0 other;

Query Match      70.0%; Score 21; DB 12; Length 1447;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1011 agacagcagggtcactccag 1031
|||||
Cp 30 AGACAGCAGCGCTCCTCCAG 10

RESULT 10
LOCUS RNP53 1627 bp RNA ROD 12-SEP-1993
DEFINITION Rat mRNA for nuclear oncoprotein p53.
ACCESSION X13058
NID 956828
KEYWORDS oncoprotein p53.
```

SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryotae; mitochondrion eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires; Rodentia;
Sciurognathi; Myomorpha; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1627)
AUTHORS Soussi, T.
TITLE Direct Submission
JOURNAL Submitted (26-SEP-1988) to the EMBL/Genbank/DBJ databases. Soussi
T., Université Pierre et Marie Curie, Unité d'Onologie
Moléculaire, IRSC - CNRS, BP 08 94802 Villejuif, France
2 (bases 1 to 1627)
AUTHORS Soussi, T., Caron de Fromental, C., Breugnot, C. and May, E.
TITLE Nucleotide sequence of a cDNA encoding the rat p53 nuclear
oncoprotein
JOURNAL Nucleic Acids Res. 16 (23), 11384 (1988)
MEDLINE 89083585
COMMENT Data kindly reviewed (09-Feb-1989) by Soussi T.
FEATURES
source
1..1627
/organism="Rattus norvegicus"
/cell_line="Py 721"
24..1199
/note="nuclear protein p53 (AA 1 - 391)"
/codon_start=1
/db_xref="PID:g56829"
/translation="MEDSOSDMSIELPLSQETFSGLKILPPDILPTATGSPNSME
DLIPDVALEEGEALVSGAPADPVTEGPAVAPATPWPJSSVPSQKTYQG
GNYGFLHGLASGTAKSVNCTYSISLAKLECOLAKTQPVQLVVSTPPESTVRMAI
YKKSQAHTEVVRBCPHHERGSDGGLAPQHLINVEGNPAEYLDIDRQTERHVVVY
EPPEVSGDYTTIHYKYNCSNGMGANRRPILITITLEDSSGNLLGDSFEVRVACAP
GRDRRTIEENFRKEKVELPELPPSGAKRALPTSTSSSPQKKRPJDEGEYTTIAIKGRK
RFEMREINLEALDKDAHTEESGDSRAHSYVTKKGGSTSRHKRMKVKVGPDSO"

BASE COUNT 374 a 486 c 380 g 381 t
ORIGIN

Query Match 70.0%; Score 21; DB 69; Length 1627;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1097 agacagcaggctcactccag 1117
|||||
Cp 30 AGACAGCAGGCTCCTCAG 10

RESULT 11
LOCUS MUSEP53M 1772 bp mRNA ROD 03-MAY-1985
DEFINITION Mouse p53 cellular tumor antigen, mRNA.
ACCESSION K01700
NID g200204
KEYWORDS DNA-binding protein; antigen; p53 gene; tumor antigen.
SOURCE Mouse embryo F9 carcinoma cells, cDNA to mRNA, clones p176, p271,
p422, p208 [1], and clone p53-1 [2].
ORGANISM Mus musculus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.

REFERENCE 1 (bases 1 to 1716)
AUTHORS Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Girol, D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306, 594-597 (1983)
MEDLINE 84068204
REFERENCE 2 (bases 105 to 1772)

AUTHORS Pennica, D., Goeddel, D.V., Hayflick, J.S., Reich, N.C., Anderson, C.M.
and Levine, A.J.
TITLE The amino acid sequence of murine p53 determined from a cDNA clone
JOURNAL Virology 134, 477-482 (1984)
MEDLINE 86072016
COMMENT The murine p53 protein is similar to the avian and human myc gene
products and the adenovirus E1a proteins. All three of these
proteins are rich in proline, contain proline runs or clusters, and
are localized in the nucleus of transformed cells. The E1a proteins
and p53 have been shown to have very short half-lives [2]. The
sequence of cDNA in [1] was established by analysing four separate
cDNA clones; p176 (bp 1-1247), p271 (bp 69-638), p422 (bp
1288-1719), and p208 (bp 1432-1719) [1].
There is only one functional p53 gene in the mouse genome. All
existing different forms of murine p53 must be products of the same
gene, mostly due to post-translational modifications [1]. The
sequence of the murine p53 pseudogene, also reported by [1] (see
separate entry), and the cDNA sequence are almost identical from
nucleotide 186 onward. Upstream of this position the two sequences
diverge totally and no homology can be observed; downstream of here
the two sequences differ by only 4%. The differences are due to
substitutions and to some small deletions or additions in this gene
relative to the cDNA [1].

FEATURES
source
1..1772
/organism="Mus musculus"
/cell_line="F9"
/dev_stage="embryo"
/sequenced_mol="cDNA to mRNA"
/tissue_type="carcinoma"
1..1668
/gene="p53"
/map="11"
158..1330
/gene="p53"
/map="11"
/codon_start=1
/product="cellular tumor antigen"
/db_xref="PID:g200205"
/translation="MTAMEESQSDISIELPLSQETFSGLKILPPDILPTSPHGMIDL
LIPQVDEEFEGSEALVSGAPADPVTEGPAVAPATPWPJSSVPSQKTYQG
NNGFHLGFLASGTAKSVNCTYSISLAKLECOLAKTQPVQLVVSTPPESTVRMAI
YKKSQAHTEVVRBCPHHERGSDGGLAPQHLINVEGNPAEYLDIDRQTERHVVVY
EPPEAGSEYTTIHYKYNCSNGMGANRRPILITITLEDSSGNLLGDSFEVRVACAPG
RDRRTIEENFRKEKVELPELPPSGAKRALPTSTSSSPQKKRPJDEGEYTTIAIKGRK
RFEMREINLEALDKDAHTEESGDSRAHSYVTKKGGSTSRHKRMKVKVGPDSO"

BASE COUNT 385 a 546 c 412 g 429 t
ORIGIN

Query Match 70.0%; Score 21; DB 67; Length 1772;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1228 agacagcaggctcactccag 1248
|||||
Cp 30 AGACAGCAGGCTCCTCAG 10

RESULT 12
LOCUS MUSEP53R 1773 bp RNA ROD 12-SEP-1993
DEFINITION Mouse mRNA for cellular tumour antigen p53.
ACCESSION X01237 K01700
NID g53575
KEYWORDS antigen; tumor antigen.

SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Givol, D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306 (5943), 594-597 (1983)
MEDLINE 84068204
REFERENCE
AUTHORS Bienz, B., Zakut-Houri, R., Givol, D. and Oren, M.
TITLE Analysis of the gene coding for the murine cellular tumour antigen p53
JOURNAL EMBO J. 3 (9), 2179-2183 (1984)
MEDLINE 85027173
REFERENCE
AUTHORS Givol, R.
TITLE Direct Submission
JOURNAL Submitted (28-NOV-1985) to the EMBL/Genbank/DBJ databases
FEATURES
source Location/Qualifiers
1..1773
/organism="Mus musculus"
158..1330
/note="p53 polypeptide (aa 1-390)"
/db_xref="PID:p53576"
/codon_start=1
/translation="MTAMEESQSDISLEPLSGPTESGLMKLPEDPILSPHGMDDL
LLQVDVEEFEGSEALRVSGAPADQVPTEPPGPAPAPATPPLSSVTPSQXTYGC
NYGFHLGFGSTAKSVMTYSPYLKLCQLVTCVQLWVSATPPASRVRYMAY
KKSQHTEVVRCPHHERCGDGLAPQHLRVENLPEYLEDROTFRHSVVVPE
PPRGESEYTTTHKYMCSGCGMRRILITITLEDSSGNLGRSEEVAVCAQC3
RDRTTEENFRKEVLCPELPSPSARALPTCTASPPQKKPJDGCEYTLTKIGRRR
FEMRELNEALELKDAAETESGSRHSSYLTKKQSTSHKKTWKVGPPSD"

old_sequence 301
/note="G is A in [1]"
/citation=[1]
old_sequence 404
/note="U is C in [1]"
/citation=[1]
old_sequence 561
/note="U is C in [1]"
/citation=[1]
old_sequence 635
/note="U is C in [1]"
/citation=[1]
old_sequence 660
/note="A is G in [1]"
/citation=[1]
old_sequence 859
/note="G is U in [1]"
/citation=[1]
old_sequence 1445..1446
/note="UA is CU in [1]"
/citation=[1]
old_sequence 1447..1448
/note="CU is UC in [1]"
/citation=[1]
old_sequence 1503
/note="U is A in [1]"
/citation=[1]
old_sequence 1558..1559

/note="CU is UC in [1]"
/citation=[1]
BASE COUNT 386 a 545 c 412 g 430 t
ORIGIN
Query Match 70.0%; Score 21; DB 67; Length 1773;
Best local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1228 agacagcagcgctcactccag 1248
|||||
Cp 30 AGACAGCAGCGCTCCTCCAG 10

RESULT 13
LOCUS MUP53P3 2130 bp DNA ROD 03-MAY-1985
DEFINITION Mouse p53 cellular tumour antigen pseudogene.
ACCESSION K02110
MID 9200206
KEYWORDS antigen; p53 gene; processed pseudogene; tumor antigen.
SOURCE Mouse 3.3-kb fragment isolated from a BALB/c genomic library, clone pch53-11.
ORGANISM Mus musculus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
1 (bases 1 to 2130)
AUTHORS Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Givol, D.
TITLE A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306, 594-597 (1983)
MEDLINE 84068204
COMMENT
The DNA sequence of pch53-11 contains a long poly-A tract, lacks
introns, and is bounded by direct repeats (bp 169-181 and bp
1852-1864), suggesting that it is a processed gene which resulted
from reverse transcription of the mature mRNA.
The sequence of the murine p53 cDNA, also reported by [1] (see
separate entry), and the pseudogene are almost identical from
nucleotide 186 onward. Upstream of this position the two sequences
diverge totally and no homology can be observed; downstream of here
the two sequences differ by only 4%. The differences are due to
substitutions and to some small deletions or additions in this gene
relative to the cDNA.
FEATURES
source Location/Qualifiers
1..2130
/organism="Mus musculus"
mRNA
/note="pseudo-p53 mRNA"
CDS
262..>1434
/note="pseudo-p53"
/pseudo
/codon_start=1
BASE COUNT 602 a 585 c 456 g 487 t
ORIGIN 19 bp upstream of BglII site.
Query Match 70.0%; Score 21; DB 67; Length 2130;
Best local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1330 agacagcagcgctcactccag 1350
|||||
Cp 30 AGACAGCAGCGCTCCTCCAG 10

RESULT 14
LOCUS MUP53P 2132 bp DNA ROD 07-NOV-1985

DEFINITION Mouse pseudogene for cellular tumour antigen p53.
ACCESSION X01236 K02110
NID 953574
KEYWORDS antigen; pseudogene; tumor antigen.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
1 (bases 1 to 2132)
REFERENCE Zakut-Houri,R., Oren,M., Bienz,B., Lavie,V., Hazum,S. and Givol,D. A single gene and a pseudogene for the cellular tumour antigen p53
JOURNAL Nature 306 (5943), 594-597 (1983)
MEDLINE 84068204
FEATURES
source Location/Qualifiers
1..2132
misc_feature
1..185 /organism="Mus musculus"
/note="5' flanking region without homology to p53 cDNA"
repeat_region
169..181 /note="direct repeat 1"
misc_feature
186..1773 /note="region of 96% homology to p53 cDNA"
misc_feature
1809..1814 /note="region of 96% homology to p53 cDNA"
polyA_site
1828 /note="pot. polyadenylation signal"
repeat_region
1854..1866 /note="pot. polyA site"
BASE COUNT 602 a 586 c 457 g 487 t
ORIGIN
Query Match 70.0%; Score 21; DB 64; Length 2132;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1331 agacagcaggctcactccag 1351
Cp 30 AGACAGCAGGCTCCTCAG 10
RESULT 15
LOCUS RNU07019 1273 bp DNA 02-FEB-1996
DEFINITION Rattus norvegicus Wistar clone p53p1 p53 pseudogene.
ACCESSION U07019
NID g460923
KEYWORDS .
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Rattus.
1 (bases 1 to 1273)
REFERENCE Lin,Y. and Chan,S.H. Cloning and characterization of two processed p53 pseudogenes from the rat genome
JOURNAL Gene 156 (2), 183-189 (1995)
MEDLINE 95278743
REFERENCE 2 (bases 1 to 1273)
AUTHORS Lin,Y.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-1994) Yue Lin, Department of Microbiology,

National University of Singapore, Lower Kent Ridge Road, Singapore 0511, Singapore
FEATURES
source Location/Qualifiers
1..1273
/clone="p53p1"
/strain="Wistar"
/organism="Rattus norvegicus"
/sex="male"
/cell_type="hepatocyte"
/tissue_type="liver"
1..14
/note="corresponds to a part of exon 1 of the rat p53 gene"
exon
/pseudo
15..97
/note="corresponds to exon 2 of the rat p53 gene"
CDS
/pseudo
24..1245
/gene="p53"
/pseudo
98..119
/gene="p53"
/note="corresponds to exon 3 of the rat p53 gene"
exon
/pseudo
120..358
/gene="p53"
/note="corresponds to exon 4 of the rat p53 gene"
exon
/pseudo
359..539
/gene="p53"
/note="corresponds to exon 5 of the rat p53 gene"
exon
/pseudo
540..640
/gene="p53"
/note="corresponds to exon 6 of the rat p53 gene"
exon
/pseudo
641..750
/gene="p53"
/note="corresponds to exon 7 of the rat p53 gene"
exon
/pseudo
751..887
/gene="p53"
/note="corresponds to exon 8 of the rat p53 gene"
exon
/pseudo
888..961
/gene="p53"
/note="corresponds to exon 9 of the rat p53 gene"
exon
/pseudo
962..1153
/gene="p53"
/note="corresponds to exon 10 of the rat p53 gene"
exon
/pseudo
1154..>1273
/note="corresponds to a part of exon 11 of the rat p53 gene"
BASE COUNT 328 a 356 c 306 g 283 t
ORIGIN
Query Match 66.7%; Score 20; DB 70; Length 1273;
Best Local Similarity 88.5%; Pred. No. 4.36e-02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Db 1039 agacagcaggctcactcctctcc 1064

Mar 25 02:50

US-08-644-289-4.rge

17

|||||
Cp 30 AGACAGCAGGGCTCCTCAGATCC 5

Search completed: Tue Mar 25 02:52:18 1997
Job time : 63 secs.

Mar 25 02:52

US-08-644-289-4.rst

1

US-08-644-289-4 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPerch_m n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Mar 25 02:53:15 1997; MasPar time 56.30 Seconds

Tabular output not generated. 234.987 Million cell updates/sec

Title: >US-08-644-289-4
Description: (1-30) from US08644289.seq
Perfect Score: 30

N.A. Sequence: 1 AGTCGGATCCTCGAGTCAGCCCTGCTCT 30
Comp: TCAGCTAGACACTCTCGGACGACAGACA

Scoring table: TABLE default
Gap 10

Mismatch STD : Dbase 0; Query 0

Searched: 630489 seqs, 220513910 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99
EST-STS-TWO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129

Mar 25 02:52

US-08-644-289-4.rst

2

130:STS2 131:STS3 132:STS4 133:STS5 134:STS6 135:STS7
136:STS8 137:STS9 138:STS10 139:STS11 140:STS12
141:STS13 142:STS14 143:STS15 144:STS16 145:STS17
146:STS18 147:STS19 148:STS20 149:STS21 150:STS22
151:STS23 152:STS24 153:STS25 154:STS26
155:STS27 156:STS28 157:STS29 158:STS30 159:STS31
160:STS32 161:STS33 162:STS34 163:STS35 164:STS36
165:STS37 166:STS38 167:STS39 168:STS40
169:STS41 170:STS42 171:STS43 172:STS44
173:STS45 174:STS46

Statistics: Mean 7.041, Variance 1.665, scale 4.228

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	ID	Description	Pred. No.
1	17	56.7	269 138	HUMUT541B	Human STS UT5411, 3'	8.05e-03		
2	17	56.7	295 75	R12972	Y770B10.r1 Homo sapie	8.05e-03		
3	17	56.7	303 46	HD0303E06B	Human fetal brain cDN	8.05e-03		
4	17	56.7	377 106	T18731	2C01B05-T3 Zea mays c	8.05e-03		
5	17	56.7	448 10	H14461	Y125F03.r1 Homo sapie	8.05e-03		
6	16	53.3	242 110	T31582	EST35170 Homo sapiens	1.00e-01		
7	16	53.3	291 4	CELK051D7R	C.elegans cDNA clone	1.00e-01		
8	16	53.3	305 147	W05720	z89B07.r1 Scores fet	1.00e-01		
9	16	53.3	305 164	HS720327	z89B07.r1 Scores fet	1.00e-01		
10	16	53.3	319 5	CELK090H9R	C.elegans cDNA clone	1.00e-01		
11	16	53.3	320 16	H32097	EST106885 Rattus sp.	1.00e-01		
12	16	53.3	324 4	CELK051E7R	C.elegans cDNA clone	1.00e-01		
13	16	53.3	327 74	R10375	YF37B06.s1 Homo sapie	1.00e-01		
14	16	53.3	329 5	CELK110H9R	C.elegans cDNA clone	1.00e-01		
15	16	53.3	334 5	CELK111C2R	C.elegans cDNA clone	1.00e-01		
16	16	53.3	337 78	R22108	YH25F03.r1 Homo sapie	1.00e-01		
17	16	53.3	338 6	CELK116G8R	C.elegans cDNA clone	1.00e-01		
18	16	53.3	383 49	HUM509A08B	Human Placenta cDNA 5	1.00e-01		
19	16	53.3	395 125	T88758	Yd90d03.r1 Homo sapie	1.00e-01		
20	16	53.3	400 173	HS129297	human STS SHGC-15280	1.00e-01		
21	16	53.3	400 134	G18129	human STS SHGC-15280	1.00e-01		
22	16	53.3	412 78	R22080	YH25B03.r1 Homo sapie	1.00e-01		
23	16	53.3	413 9	H08944	Y193e07.r1 Homo sapie	1.00e-01		
24	16	53.3	476 148	W07011	z91F04.r1 Scores fet	1.00e-01		
25	16	53.3	476 157	HS011323	z91F04.r1 Scores fet	1.00e-01		
26	16	53.3	487 89	R60515	YH13e11.r1 Homo sapie	1.00e-01		
27	16	53.3	515 120	T70205	Yc18F04.s1 Homo sapie	1.00e-01		
28	16	53.3	530 49	HUM510G08B	Human Placenta cDNA 5	1.00e-01		
29	16	53.3	537 36	H97840	Yw09C03.s1 Homo sapie	1.00e-01		
30	15	50.0	208 71	OSU37917	Oryza sativa clone pf	1.13e+00		
31	15	50.0	217 70	N71886	z90F03.s1 Homo sapie	1.13e+00		
32	15	50.0	248 40	HSC0YC012	H. sapiens partial cd	1.13e+00		
33	15	50.0	333 174	HSB0542E9	H. sapiens (D17S1839)	1.13e+00		
34	15	50.0	341 133	G14170	human STS SHGC-6733 c	1.13e+00		
35	15	50.0	345 35	H93667	Yw55C12.r1 Homo sapie	1.13e+00		
36	15	50.0	382 115	T52077	Yb10e11.r1 Homo sapie	1.13e+00		
37	15	50.0	388 160	HS4294289	Yv76E08.s1 Homo sapie	1.13e+00		
38	15	50.0	399 81	R33980	Yh74e05.r1 Homo sapie	1.13e+00		
39	15	50.0	400 134	G17772	human STS SHGC-13700	1.13e+00		
40	15	50.0	400 174	HS582301	human STS SHGC-10638	1.13e+00		
41	15	50.0	402 84	R44555	Yg29e04.s1 Homo sapie	1.13e+00		
42	15	50.0	443 90	R66640	Y135a08.r1 Homo sapie	1.13e+00		

Mar 25 02:52

US-08-644-289-4.rst

3

43 15 50.0 445 68 N70695 za32d02.s1 Homo sapie 1.13e+00
c 44 15 50.0 496 57 N34722 yx82b12.r1 Homo sapie 1.13e+00
45 15 50.0 498 13 H23094 ym51d09.s1 Homo sapie 1.13e+00

ALIGNMENTS

RESULT 1 HUM05411B 269 bp DNA STS 28-DEC-1994

LOCUS Human STS UT5411, 3' primer bind.

DEFINITION L30902

ACCESSION 9604971

KEYWORDS PCR primer; STS sequence; microsatellite DNA; microsatellite marker; sequence tagged site; tetranucleotide repeat.

SOURCE Homo sapiens

ORGANISM Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homidae.

REFERENCE 1 (bases 1 to 269)

Gerken, S.C., Matsunami, N., Plaetke, R., Albertsen, H., Ballard, L., Wells, R., Lawrence, E., Moore, M., Holik, P.R., Carlson, M., Zhao, X., Robertson, M., Bradley, P., Elener, T., Tingey, A., Lalouel, J.-M., and White, R.

Genetic and physical mapping of simple sequence repeat containing sequence tagged sites from the human genome

Unpublished (1994) See COMMENT for author address

Submitted by: Utah Center for Human Genome Research University of Utah, Dept. of Human Genetics

2160 Eccles Institute of Human Genetics

Salt Lake City, UT 84112

e-mail: sts@corona.med.utah.edu

Primer A: CAGCAGATGCTGACCTTTGGGAC

Primer B: CATTCACCCCTCAGACCTTACA

End to Label: Primer B

PCR Profile:

Initial Denaturation: 94C 300sec Extension 30 94

Cycles Denaturation Annealing C 10 sec. 74 C 10 sec. 72 C 20 sec. 30

74 C 10 sec. 72 C 20 sec. Mg++: 1.00 mM

Gel: Acrylamide 7%, Formamide 32%, Urea 34%

Alleles: 1.

FEATURES location/Qualifiers

source 1..269

/organism="Homo sapiens"

/sequenced_mol="DNA"

primer_bind complement(148..171)

/map="8"

/evidence=experimental

Mar 25 02:52

US-08-644-289-4.rst

4

ACCESSION R12972
NID 9766048
KEYWORDS EST.
SOURCE human clone=27554 library=Soares infant brain INIB vector=lafmid BA host=DH10B (ampicillin resistant) primer=M13R1 Rstrel=Not I Rstrel2=Hind III Whole brain from a 73 days post natal female. 1st strand cDNA was primed with a Not I - oligo(dT) primer 15' AACTGAGACATTCGGCGGCGAGCAATTTTCTTTT 3'; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 295)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marras, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

COMMENT GDB: G00-399-901

Contact: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

High quality sequence stops: 233

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

FEATURES location/Qualifiers

source 1..295

/organism="Homo sapiens"

/clone="27554"

/note="human"

BASE COUNT 78 a 72 c 59 g 85 t 1 others

ORIGIN

Query Match 56.7%; Score 17; DB 75; Length 295;

Best Local Similarity 83.3%; Pred. No. 8.05e-03;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 262 acagaaggtcactccatcc 285

Cp 28 ACAGACGCGCTCAGCTCAGATCC 5

RESULT 3 HUM030E0B 303 bp mRNA EST 26-AUG-1995

LOCUS Human fetal brain cDNA 5'-end GEN-030E06.

4

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
				Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata; Vertebrata; Sarcophagidii; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
				1 (bases 1 to 303)
	Fujimura,T., Hirano,H., Katagiri,T., Kamai,A., Kuga,Y., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinozuka,H., Takachi,A., Takeda,S., Watanabe,T., Takahashi,E.-I., Hirai,Y., Maekawa,H., Shin,S. and Nakamura,Y.	Unpublished (101)	Unpublished (1995)	
				Submitted (30-May-1995) to DDBJ by:
				Tsutomu Fujimura
				Otsuka GEN Research Institute
				Otsuka Pharmaceutical Co., Ltd
				463-10 Kagasuno Kawauchi-cho
				Tokushima, Tokushima
				771-01
				Japan
				Phone: 0886-65-2888
				Fax : 0886-37-1035.
FEATURES				
source				Location/Qualifiers
				1..303
				/organism="Homo sapiens"
				/sequenced_mol="cDNA to mRNA"
				/clone_11p="Clontech human fetal brain polyA+ mRNA (#6353)"
BASE COUNT	53 a	91 c	82 g	55 t 22 others
ORIGIN				
Query Match				56.7%; Score 17; DB 46; Length 303;
Best Local Similarity				62.1%; Pred. No. 8.05e-03;
Matches	18; Conservative	5; Mismatches	6; Indels	0; Gaps 0;
Db	204 agacgcacgtctctdbcbctcatccacc 232			11::: : : 11
Cp	30 AGACACGACGCGCTCAGCTCCAGATCCGAC 2			
RESULT	4			
LOCUS	118731	377 bp	mRNA	EST 12-MAY-1994
DEFINITION	2C01B05-13 Zea mays cDNA clone 2C01B05 3' end similar to 19KD alpha zein.			
ACCESSION	118731			
NID	9485661			
KEYWORDS	EST.			
SOURCE	maize clone=2C01B05 library=membrane-free polysomes from endosperm strain=M64A2 vector=Ziplox host=BD110B primer=RT181=SalI			
	Rs1=NotI ds-cDNA was prepared from oligo-dT selected mRNA by			
	priming with a NotI oligo-dT oligomer and then adding the second			
	strand to RNease-nicked DNA:RNA hybrid with DNA PolI. SalI adaptors			
	were added to the ends, the ds-cDNAs were then digested with NotI			
	and size-selected. These were directionally-cloned into the Ziplox			
	phage vector, excised as plaemids, and then individually analyzed.			
ORGANISM	Zea mays			
	Eucaryotae; Embryophyta; Magnoliophyta; Liliopsida; Cyperales;			
	Poaceae; Zea.			
REFERENCE	1 (bases 1 to 377)			
AUTHORS	Shen,B., Carmelito,N., Torres-Jerez,I., Stevenson,R., Helentjaire,T., Hadden,J., Larikne,B., Almiria,E., Ferl,R. and Baydortier,C.			
TITLE	Single-Pass Sequencing and mapping of clones from two maize cDNA libraries			
JOURNAL	Unpublished (1994)			

COMMENT	CONTACT: THE MAIZE cDNA PROJECT.
FEATURES	Location/Qualifiers
source	1..377
	/organism="Zea mays"
	/clone="ZC01B05"
	/strain="M64A2"
	/note="maize"
BASE COUNT	103 a 59 c 120 g 95 t
ORIGIN	
Query Match	56.7%; Score 17; DB 106; Length 377;
Best Local Similarity	90.5%; Pred. No. 8.05e-03;
Matches	19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db	190 aggcagcagggccacaccag 210
Cp	30 ACAGACGACGGCTCACTCCAG 10
RESULT	5
LOCUS	H14461 448 bp mRNA EST 27-JUN-1995
DEFINITION	Y125F03.c1 Homo sapiens cDNA clone 159293 5' similar to contains
ACCESSION	MER22 repetitive element ;.
NID	H14461
KEYWORDS	9879281
SOURCE	EST.
	human clone=159293 library=Soares breast 3NbHbst vector=pT7T3D
	(Pharmacita) with a modified polylinker host=DH10B (ampicillin
	resistant) primer=MD13RP1 Rsite=Not I Rsite2=Eco RI Adult human.
	1st strand cDNA was primed with a Not I - oligo(dT) primer [5'
	TGTACACATCTGAAGTCGGAGCGGCCGCTTTTCTTTTCTTTTCTTTT 3'],
	double-stranded cDNA was ligated to Eco RI adaptors (Pharmacita),
	digested with Not I and cloned into the Not I and Eco RI sites of a
	modified pT7T3 vector (Pharmacita). Library went through one round
	of normalization to a Cot = 20. Library constructed by Bento Soares
	and M.Falina Bonaldo.
ORGANISM	Homo sapiens
	Eukaryotae; Metazoa; Emmetazoa; Bilateria; Coelomata;
	Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
	Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
	Eutheria; Archonta; Primates; Catarrhini; Homiidae; Homo.
REFERENCE	1 (bases 1 to 448)
AUTHORS	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawking,M.,
	Holman,M., Holtman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
	Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
	Treaskie,E., Waterston,R., Williamson,A., Wohlmann,P. and
	Wilson,R.
TITLE	The WashU-Merck EST Project
JOURNAL	Unpublished (1995)
COMMENT	
	Contact: Wilson RK
	WashU-Merck EST Project
	Washington University School of Medicine
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
	Tel: 314 286 1800
	Fax: 314 286 1810
	Email: est@wustl.wustl.edu
	High quality sequence stops: 310
	Source: IMAGE Consortium, LNL
	This clone is available royalty-free through LNL; contact the
	IMAGE Consortium (info@image.lnl.gov) for further information.
FEATURES	Location/Qualifiers
source	1..448

Mar 25 02:52

US-08-644-289-4.rst

7

/organism="Homo sapiens"
/clone="159293"BASE COUNT 90 a 124 c 113 g 109 t 12 others
ORIGINQuery Match 56.7%; Score 17; DB 10; Length 448;
Best Local Similarity 83.3%; Pred. No. 8.05e-03;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;Db 152 gatcctggatgagccctgcgtc 175
Yy 6 GATCCTGGAGTGAAGCCCTGCTGTC 29RESULT 6
LOCUS T31582 242 bp mRNA EST 06-SEP-1995
DEFINITION EST15170 Homo sapiens cDNA 5' end similar to None..
ACCESSION T31582
NID 6133680
KEYWORDS EST.
SOURCE human primer=M13 Reverse library=Human Embryo.
ORGANISM Homo sapiensREFERENCE
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W., Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.M., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,J.P., Kelley,J.M., Kline,K.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,J.M., Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bedarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.-J., Dinke,D., Feng,P., Ferrie,A., Fischer,C., Hastings,G.A., He,M.-W., Hu,J.-S., Greene,J.M., Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H., Meisner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M., and Venter,J.C.TITLE Initial Assessment of Human Gene Diversity and Expression Patterns
Based Upon 52 Million Basepairs of cDNA Sequence
JOURNAL Unpublished (1995)
COMMENT Other ESTs: THC11831
Contact: Venter, JC
The Institute for Genomic Research
932 Clopper Rd, Gaithersburg, MD 20878
Tel: 3018699056
Fax: 3018699423
Email: tdbinfo@db.tigr.org
For clone availability, additional sequence and expression information related to this EST, please contact the TIGR Database (tdbinfo@db.tigr.org).FEATURES
source location/Qualifiers
1..242/organism="Homo sapiens"
/note="human"
RNA <1..>242
BASE COUNT 50 a 77 c 58 g 54 t 3 others

Mar 25 02:52

US-08-644-289-4.rst

8

ORIGIN

Query Match 53.3%; Score 16; DB 110; Length 242;
Best Local Similarity 90.0%; Pred. No. 1.00e-01;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;Db 22 gatcctggatgagccctgc 41
Yy 6 GATCCTGGAGTGAAGCCCTGCTG 25RESULT 7
LOCUS CELK051D7R 291 bp mRNA EST 09-DEC-1995
DEFINITION C.elegans cDNA clone yk51d7 : 3' end, single read.
ACCESSION D70927
NID g110634
KEYWORDS EST(expressed sequence tag).
SOURCE Caenorhabditis elegans (strain N2,) (library: Yuji Kohara unpublished cDNA) Hermaphrodite embryo cDNA to mRNA.
ORGANISM Caenorhabditis elegansREFERENCE
AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Sugimoto,A., Matanabe,H. and Nishigaki,A.
TITLE Toward an expression map of the C.elegans genome
JOURNAL Unpublished (1995)
COMMENT Submitted (23-Aug-1995) to DDBJ by:
Yuji Kohara
Gene Library Lab.
National Institute of Genetics
Yata 1111, Mishima Shizuoka
411 Japan
Phone: 0559-81-6854
Fax : 0559-81-6855
Email:ykohara@dbj.nig.ac.jp.FEATURES
source location/Qualifiers
1..291/organism="Caenorhabditis elegans"
/strain="N2"
/dev stage="embryo"
/sequenced_mol="cDNA to mRNA"
/sex="Hermaphrodite"
/tissue type="embryo"
/clone lib="Yuji Kohara unpublished cDNA"
BASE COUNT 88 a 55 c 67 g 80 t 1 others
ORIGINQuery Match 53.3%; Score 16; DB 4; Length 291;
Best Local Similarity 80.8%; Pred. No. 1.00e-01;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;Db 114 gtgcgactcctgagtcacatccagctg 199
Yy 2 GTGCAGTCTCGAGTGAAGCCCTGCTG 27RESULT 8
LOCUS W05720 305 bp mRNA EST 23-APR-1996
DEFINITION za89b07.r1 Soares fetal lung NBHL19W Homo sapiens cDNA clone 299701
ACCESSION W05720
NID g1278511
5' similar to gb:U19067 TRANSCRIPTION FACTOR P65 (HUMAN);.

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 305)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Treviski, E., Waterston, R., Williamson, A., Wohlmann, P., and
Wilson, R.
TITLE The MashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
MashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watscn.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Seq primer: mob.REGA+ET
High quality sequence stop: 1.

NCBI gi: 1278511

FEATURES
source location/Qualifiers

1..305
/organism="Homo sapiens"
/note="Organ: Lung; Vector: p7T73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer:
[5'-TGTTACCATCTGACATGCGACGCCGCAATTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p7T73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NBHL19W."
/clone="299701"
/clone_lib="Soares fetal lung NBHL19W"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1..>305
BASE COUNT 81 a 66 c 82 g 66 t 10 others
ORIGIN

Query Match 53.3%; Score 16; DB 147; Length 305;
Best Local Similarity 79.2%; Pred. No. 1.00e-01;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Db 184 ggaactcgnagagcagctgcgt 207
||||| ||| ||| ||| |||
Qy 5 GGAATCTGAGTGAAGCCCTGCTGT 28

RESULT 9
ID HS720327 standard; RNA; EST; 305 BP.
AC W05720;
DT 08-MAY-1996 (Rel. 47, Created)

DT 08-MAY-1996 (Rel. 47, Last updated, Version 1)
DE za89b07.r1 Soares fetal lung NBHL19W Homo sapiens cDNA clone 299701
DE 5' similar to gb:U19067 TRANSCRIPTION FACTOR P65 (HUMAN);.
KW EST.
OS Homo sapiens (human)
OC Eukaryotae; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominiidae.
RN 11
RP 1-305
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Treviski E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RT "The MashU-Merck EST Project";
RL Unpublished.
CC Contact: Wilson RK MashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@watscn.wustl.edu This clone is available royalty-free through
CC LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for
CC further information. Trace considered overall poor quality Seq
CC primer: mob.REGA+ET High quality sequence stop: 1. NCBI gi: 1278511
FH Key location/Qualifiers
FH
FH source 1..305
/organism="Homo sapiens"
/note="Organ: lung; Vector: p7T73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer:
[5'-TGTTACCATCTGACATGCGACGCCGCAATTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p7T73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NBHL19W."
/clone="299701"
/clone_lib="Soares fetal lung NBHL19W"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1..>305
FT mRNA
FT Sequence 305 BP; 81 A; 66 C; 82 G; 66 T; 10 other;
SQ

Query Match 53.3%; Score 16; DB 164; Length 305;
Best Local Similarity 79.2%; Pred. No. 1.00e-01;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Db 184 ggaactcgnagagcagctgcgt 207
||||| ||| ||| ||| |||
Qy 5 GGAATCTGAGTGAAGCCCTGCTGT 28

RESULT 10
LOCUS CELK09049R 319 bp mRNA EST 10-DEC-1995
DEFINITION C.elegans cDNA clone yk90h9 : 3' end, single read.
ACCESSION D72177
NID g1111884
KEYWORDS EST(expressed sequence tag).
SOURCE Caenorhabditis elegans (strain N2,) (library: Yoji Kohara
unpublished cDNA) Hermaphrodite embryo cDNA to mRNA.
ORGANISM Caenorhabditis elegans
Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;

Secernentea; Rhabditia; Rhabditidae; Rhabditina; Rhabditoidea;
Rhabditidae; Caenorhabditis.

REFERENCE 1 (bases 1 to 319)

AUTHORS Kohara, Y., Mochizashi, T., Tabara, H., Sugimoto, A., Watanabe, H. and Nishigaki, A.

TITLE Toward an expression map of the *C. elegans* genome

JOURNAL Unpublished (1995)

COMMENT Submitted (23-Aug-1995) to DDBJ by:

Yuji Kohara

Gene Library Lab.

National Institute of Genetics

Yata 1111, Mishima Shizuoka

411 Japan

Phone: 0559-81-6854

Fax : 0559-81-6855

Email: ykohara@dbj.nig.ac.jp.

location/Qualifiers

FEATURES

source

1..319

/organism="Caenorhabditis elegans"

/strain="N2"

/dev_stage="embryo"

/sequenced_mol="cDNA to mRNA"

/sex="Hermaphrodite"

/tissue_type="embryo"

BASE COUNT /clone_lib="Yuji Kohara unpublished cDNA"

ORIGIN 94 a 57 c 80 g 88 t

Query Match 53.3%; Score 16; DB 5; Length 319;

Best Local Similarity 80.8%; Pred. No. 1.00e-01;

Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 164 gtcgagtcctgagtcacatccagctg 189

0y 2 gtcgagtcctgagtcacatccagctg 27

RESULT 11

LOCUS H32097 320 bp mRNA EST 08-SEP-1995

DEFINITION EST106885 Rattus sp. cDNA 5' end similar to Cadherin 43.

ACCESSION H32097

NID g977514

KEYWORDS EST.

SOURCE rat primer-M13 Reverse library=Rat PC-12 cells, untreated

vector=pluscript SK- Ratel=EcotRI Ratel2=XhoI poly(A)+ RNA was
purified from untreated PC12 cells cultured for 9 days. cDNA was
constructed using an oligo-dT primer and directionally cloned using
the Lambda Zap II Vector Kit by Stratagene.

ORGANISM Rattus sp.

Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Myomorpha;

Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 320)

AUTHORS Lee, N.H., Weinstein, K.G., Kirtress, E.F., Earle-Hughes, J.A.,

Fulmer, R.A., Marmaraz, S., Glodok, A., Gocayne, J.D., Adams, M.D.,

Kerlavage, A.R., Frazer, C.M. and Venter, J.C.

TITLE Comparative expressed sequence tag analysis of differential gene

expression profiles in PC-12 cells before and after nerve growth

factor treatment

JOURNAL Proc. Natl. Acad. Sci. USA 92, 8303-8307 (1995)

COMMENT

Contact: Lee NH

The Institute for Genomic Research

932 Clopper Rd, Gaithersburg, MD 20878

Tel: 3018699056

Fax: 3018699423

Email: nhlee@tigr.org

For clone availability please contact the TIGR Database

(tldbinfo@tldb.tigr.org).

location/Qualifiers

1..320

/organism="Rattus sp."

/note="rat"

<1..>320

BASE COUNT 85 a 98 c 86 g 49 t 2 others

ORIGIN mRNA

Query Match 53.3%; Score 16; DB 16; Length 320;

Best Local Similarity 94.4%; Pred. No. 1.00e-01;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 246 cagggtcaccaccagat 263

Cp 24 CAGGCTCATTCCAGGAT 7

RESULT 12

LOCUS CELK051E/R 324 bp mRNA EST 09-DEC-1995

DEFINITION C.elegans cDNA clone yK51e7 : 3' end, single read.

ACCESSION D70933

NID q1110642

KEYWORDS EST(expressed sequence tag).

SOURCE Caenorhabditis elegans (strain N2,) (library: Yuji Kohara

unpublished cDNA) Hermaphrodite embryo embryo cDNA to mRNA.

ORGANISM Caenorhabditis elegans

Eukaryote; mitochondrial eukaryotes; Metazoa; Nematoda;

Secernentea; Rhabditia; Rhabditidae; Rhabditina; Rhabditoidea;

Rhabditidae; Caenorhabditis.

REFERENCE 1 (bases 1 to 324)

AUTHORS Kohara, Y., Mochizashi, T., Tabara, H., Sugimoto, A., Watanabe, H. and

Nishigaki, A.

TITLE Toward an expression map of the *C. elegans* genome

JOURNAL Unpublished (1995)

COMMENT Submitted (23-Aug-1995) to DDBJ by:

Yuji Kohara

Gene Library Lab.

National Institute of Genetics

Yata 1111, Mishima Shizuoka

411 Japan

Phone: 0559-81-6854

Fax : 0559-81-6855

Email: ykohara@dbj.nig.ac.jp.

location/Qualifiers

1..324

/organism="Caenorhabditis elegans"

/strain="N2"

/dev_stage="embryo"

/sequenced_mol="cDNA to mRNA"

/sex="Hermaphrodite"

/tissue_type="embryo"

/clone_lib="Yuji Kohara unpublished cDNA"

BASE COUNT 94 a 58 c 81 g 89 t 2 others

ORIGIN

Query Match 53.3%; Score 16; DB 4; Length 324;

Best Local Similarity 80.8%; Pred. No. 1.00e-01;

Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 170 gtcgacatcctgagtcacatccagctg 195
|||||
Oy 2 CTCGACATCTGAGTACGCCCTGCTG 27

RESULT 13
LOCUS RI0375 327 bp mRNA EST 06-APR-1995
DEFINITION y137b06.s1 Homo sapiens cDNA clone 129011 3'.
ACCESSION RI0375
NID g162331
KEYWORDS EST.
SOURCE human clone=129011 library=Soares fetal liver spleen INFIS
vector=pf773D (Pharmacia) with a modified polylinker host=DH10B
(ampicillin resistant) primers=21m13 Rsite1-Pac I Rsite2-Eco RI
liver and spleen from a 20 week-post conception male fetus. 1st
strand cDNA was primed with a Pac I - oligo (dT) primer [5'
AACTGCAAGATTAATTAAGATCTTTTCTTTTCTTTTCTTTT 3'], double-stranded
cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
I and cloned into the Pac I and Eco RI sites of the modified p773
vector. Library went through one round of normalization. Library
constructed by Bento Soares and M. Patricia Bonaldo.

ORGANISM Homo sapiens
Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 327)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Maier, M.,
Parsavani, V., Rafkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P., and
Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 163
Source: IMAGE Consortium, LNL.
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
LOCATION/Qualifiers
1..327
/organism="Homo sapiens"
/clone="129011"
/note="human"

BASE COUNT 69 a 88 c 79 g 81 t 10 others
ORIGIN

Query Match 53.3%; Score 16; DB 74; Length 327;
Best Local Similarity 94.4%; Pred. No. 1.00e-01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 229 ctgagatgagccacagctg 246
|||||
Oy 10 CTCGACGACGCCCTGCTG 27

RESULT 14
LOCUS CELK1101HR 329 bp mRNA EST 10-DEC-1995
DEFINITION C.elegans cDNA clone yk110h11 : 3' end, single read.

ACCESSION D72892
NID g1112607
KEYWORDS EST(expressed sequence tag).
SOURCE Caenorhabditis elegans (strain N2,) (library: Yuji Kohara
unpublished cDNA) Hermaphrodite embryo cDNA to mRNA.
ORGANISM Caenorhabditis elegans
Eucaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabdilia; Rhabdilitidae; Rhabdilitina; Rhabdilitoidea;
Rhabdilitidae; Caenorhabditis.
REFERENCE 1 (bases 1 to 329)
AUTHORS Kohara, Y., Motobash, T., Tabara, H., Sugimoto, A., Watanabe, H. and
Nishigaki, A.
TITLE Toward an expression map of the C.elegans genome
JOURNAL Unpublished (1995)
COMMENT Submitted (23-Aug-1995) to DDBJ by:
Yuji Kohara
Gene Library Lab.
National Institute of Genetics
Yata 1111, Mishima Shizuoka
411 Japan
Phone: 0559-81-6854
Fax : 0559-81-6855
Email:ykohara@dbj.nig.ac.jp.

FEATURES
source
1..329
/organism="Caenorhabditis elegans"
/strain="N2"
/dev_stage="embryo"
/sequenced_mol="cDNA to mRNA"
/sex="Hermaphrodite"
/tissue_type="embryo"
/clone_lib="Yuji Kohara unpublished cDNA"

BASE COUNT 95 a 57 c 83 g 90 t 4 others
ORIGIN

Query Match 53.3%; Score 16; DB 5; Length 329;
Best Local Similarity 80.8%; Pred. No. 1.00e-01;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 174 gtcgacatcctgagtcacatccagctg 199
|||||
Oy 2 CTCGACATCTGAGTACGCCCTGCTG 27

RESULT 15
LOCUS CELK111CZR 334 bp mRNA EST 10-DEC-1995
DEFINITION C.elegans cDNA clone yk111c12 : 3' end, single read.
ACCESSION D72916
NID g1112628
KEYWORDS EST(expressed sequence tag).
SOURCE Caenorhabditis elegans (strain N2,) (library: Yuji Kohara
unpublished cDNA) Hermaphrodite embryo cDNA to mRNA.
ORGANISM Caenorhabditis elegans
Eucaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabdilia; Rhabdilitidae; Rhabdilitina; Rhabdilitoidea;
Rhabdilitidae; Caenorhabditis.
REFERENCE 1 (bases 1 to 334)
AUTHORS Kohara, Y., Motobash, T., Tabara, H., Sugimoto, A., Watanabe, H. and
Nishigaki, A.
TITLE Toward an expression map of the C.elegans genome
JOURNAL Unpublished (1995)
COMMENT Submitted (23-Aug-1995) to DDBJ by:
Yuji Kohara
Gene Library Lab.

Mar 25 02:52

US-08-644-289-4.fst

15

National Institute of Genetics
Yata 1111, Mishima Shizuoka

411 Japan

Phone: 0559-81-6854

Fax : 0559-81-6855

Email: ykohara@dbj.nig.ac.jp.

location/Qualifiers

source

1..334

/organism="Caenorhabditis elegans"

/strain="N2"

/dev_stage="embryo"

/sequenced_mol="cDNA to mRNA"

/sex="Hermaphrodite"

/tissue_type="embryo"

/clone_lib="Yuji Kohara unpublished cDNA"

BASE COUNT 97 a 60 c 86 g 90 t 1 others

ORIGIN

Query Match

Best Local Similarity 80.8%; Score 16; DB 5; Length 334;

Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 169 gtcgatacctgagtcacacgctg 194

|||||

Qy 2 gtcgatacctgagtcacacgctg 27

Search completed: Tue Mar 25 02:54:20 1997
Job time : 65 secs.

Mar 25 02:55

US-08-644-289-5.lgc

1

Mar 25 02:55

US-08-644-289-5.lgc

2

INTELLECTUAL
PROPERTY
(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPearch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Mar 25 02:56:08 1997; Maspark time 49.48 Seconds

166,449 Million cell updates/sec

Tabular output not generated.

Title: >US-08-644-289-5

Description: (1-10) from US08644289.seq

Perfect Score: 10

N.A. Sequence: 1 AGCGATGCGT 10

Comp: TCCGTACGCA

Scoring table: TABLE default

Gap 10

Mismatch STD : Dbase 0; Query 0

Searched: 279077 seqs, 41180865 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb1-new5

1:BCT 2:FUN 3:INV1 4:INV2 5:ORG

10:PRO1 11:PRO2 12:ROD 13:SYN 14

Database: genbank94

16:BCT1 17:BCT2 18:BCT3 19:BCT4

23:BCT8 24:BCT9 25:INV1 26:INV2 27:INV3 28:INV4 29:INV5 30:INV6 31:INV7 32:MAM1 33:MAM2 34:MAM3 35:VRL1 36:VRL2

37:VRL3 38:PAT1 39:PAT2 40:PAT3 41:PHG 42:PLN1 43:PLN2

44:PLN3 45:PLN4 46:PLN5 47:PLN6 48:PLN7 49:PLN8 50:PR11

51:PR12 52:PR13 53:PR14 54:PR15 55:PR16 56:PR17 57:PR18

58:PR19 59:PR110 60:PR111 61:PR112 62:PR113 63:ROD1

64:ROD2 65:ROD3 66:ROD4 67:ROD5 68:ROD6 69:ROD7 70:ROD8

71:STR 72:SYN 73:UNA 74:VRL1 75:VRL2 76:VRL3 77:VRL4

78:VRL5 79:VRL6 80:VRL7 81:VRL8

Database: genbank-new5

82:BCT 83:INV1 84:INV2 85:MAM 86:VRL 87:PAT 88:PHG

89:PLN 90:PRI 91:ROD 92:STR 93:SYN 94:UNA 95:VRL

Database: u-emb146_94

96:part1

Statistics: Mean 5.503; Variance 1.579; scale 3.486

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Ex. Eyer-
Sag ID 75 had many 100%
matches. If it would be helpful,
more gene alignments can be
printed, or seqs retrieved can be
limited by length.
Please call if you want this
done. Barb Olsen 308-4291

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	10	100.0	153	66	MUSCC210	Mouse complement comp	8.12e+02
2	10	100.0	427	84	EGHBX2	E.granulosus Eghbx2 h	8.12e+02
3	10	100.0	956	82	MEU41091	Unidentified marine e	8.12e+02
4	10	100.0	956	82	MEU41088	Unidentified marine e	8.12e+02
5	10	100.0	2060	90	HOMIGMT	Homo sapiens beta-1,6	8.12e+02
6	10	100.0	2613	90	HSSP3A120	H.sapiens mRNA for sp	8.12e+02
7	10	100.0	3225	96	S74971	STM-catecholamine-eul	8.12e+02
8	10	100.0	3450	83	DROADHBO39	Drosophila pseudobsc	8.12e+02
9	10	100.0	3463	83	DROADHBO32	Drosophila pseudobsc	8.12e+02
10	10	100.0	3465	83	DROADHBO35	Drosophila pseudobsc	8.12e+02
11	10	100.0	3466	83	DROADHBO31	Drosophila pseudobsc	8.12e+02
12	10	100.0	3471	83	DROADHBO32	Drosophila pseudobsc	8.12e+02
13	10	100.0	3472	83	DROADHBO33	Drosophila pseudobsc	8.12e+02
14	10	100.0	3472	83	DROADHBO34	Drosophila pseudobsc	8.12e+02
15	10	100.0	3472	83	DROADHBO35	Drosophila pseudobsc	8.12e+02
16	10	100.0	3475	83	DROADHBO32	Drosophila pseudobsc	8.12e+02
17	10	100.0	3475	83	DROADHBO33	Drosophila pseudobsc	8.12e+02
18	10	100.0	3475	83	DROADHBO34	Drosophila pseudobsc	8.12e+02
19	10	100.0	3475	83	DROADHBO35	Drosophila pseudobsc	8.12e+02
20	10	100.0	3476	83	DROADHBO32	Drosophila pseudobsc	8.12e+02
21	10	100.0	3479	83	DROADHBO33	Drosophila pseudobsc	8.12e+02
22	10	100.0	3481	83	DROADHBO34	Drosophila pseudobsc	8.12e+02
23	10	100.0	3485	83	DROADHBO35	Drosophila pseudobsc	8.12e+02
24	10	100.0	3487	83	DROADHBO32	Drosophila pseudobsc	8.12e+02
25	10	100.0	3487	83	DROADHBO33	Drosophila pseudobsc	8.12e+02
26	10	100.0	3494	83	DROADHBO34	Drosophila pseudobsc	8.12e+02
27	10	100.0	3506	83	DROADHBO35	Drosophila pseudobsc	8.12e+02
28	10	100.0	3827	89	SCYNL243W	S.cerevisiae chromoso	8.12e+02
29	10	100.0	5919	89	SCYNL242W	S.cerevisiae chromoso	8.12e+02
30	10	100.0	6508	82	PSECUMA	Pseudomonas fluoresce	8.12e+02
31	10	100.0	6508	82	PSECUMA	Pseudomonas fluoresce	8.12e+02
32	10	100.0	10019	82	HU32716	Haemophilus influenza	8.12e+02
33	10	100.0	11473	67	MUSPCADH2	Mouse P-cadherin gene	8.12e+02
34	10	100.0	12394	82	HU32842	Haemophilus influenza	8.12e+02
35	10	100.0	19371	83	CEK10C9	Caenorhabditis elegans	8.12e+02
36	10	100.0	30061	83	CEC30D11	Caenorhabditis elegans	8.12e+02
37	10	100.0	34842	84	CEP54C9	Caenorhabditis elegans	8.12e+02
38	10	100.0	35743	83	CEP31E8	Caenorhabditis elegans	8.12e+02
39	10	100.0	39496	84	CEP13D12	Caenorhabditis elegans	8.12e+02
40	10	100.0	40867	83	CEP26F12	Caenorhabditis elegans	8.12e+02
41	10	100.0	41375	90	HSN27C5	Human DNA sequence fr	8.12e+02
42	10	100.0	42503	90	HSU15G11	Human DNA sequence fr	8.12e+02
43	10	100.0	125787	90	HS13817A	Human DNA sequence fr	8.12e+02
44	10	100.0	282700	82	BACCH642	Bacillus subtilis DNA	8.12e+02
45	10	100.0	282700	82	BACCH642	Bacillus subtilis DNA	8.12e+02

ALIGNMENTS

RESULT	1	MUSCC210	153 bp	DNA	ROD	18-JAN-1991
LOCUS						
DEFINITION						
ACCESSION						
NID						
KEYWORDS						
SEGMENT						
SOURCE						
ORGANISM						

Mar 25 02:55

US-08-644-289-5.lge

3

Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE 1 (bases 1 to 153)
AUTHORS Iehikawa,N., Nonaka,M., Wetsel,R.A. and Colten,H.R.
TITLE Murine complement C2 and factor B genomic and cDNA cloning reveals
JOURNAL different mechanisms for multiple transcripts of C2 and B
MEDLINE J. Biol. Chem. 265, 19040-19046 (1990)
91035430

FEATURES
source Location/Qualifiers
1..153
/organism="Mus musculus"
/sequenced_mol="DNA"
order(M60571:97..102,1..6)
/gene="C2"
/note="2.1 kb gap"
/number=9
exon 7..147
/gene="C2"
/number=10
BASE COUNT 40 a 28 c 50 g 35 t
ORIGIN

Query Match 100.0%; Score 10; DB 66; Length 153;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 87 aggcattgcct 96
|||||
Qy 1 AGGCATGCCT 10

RESULT 2
LOCUS EGHX2 427 bp DNA INV 23-APR-1996
DEFINITION E.granulosus Eghx2 homeobox gene.
ACCESSION X66818
NID 91279419
KEYWORDS homeobox gene.
SOURCE Echinococcus granulosus.
ORGANISM Echinococcus granulosus.
REFERENCE 1 (bases 1 to 427)
AUTHORS Oliver,G.C.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1992) G.C. Oliver, Max-Planck-Inst of Biophysical
REMARK Chemistry, Dept of Cell Biology, Am Fassberg 3400, Goettingen, FRC
Revised by [3]
2 (bases 1 to 180)
AUTHORS Oliver,G., Vispo,M., Mailhos,A., Martinez,C., Sosa-Pineda,B.,
Fieletz,W. and Ehrlich,R.
TITLE Homeoboxes in flatworms
JOURNAL Gene 121 (2), 337-342 (1992)
MEDLINE 93077050
REFERENCE 3 (bases 1 to 427)
AUTHORS Vispo,M.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1996) Vispo M., Dpto. Genetica. Facultad de
Biologia, Av. Diagonal 645, 08071 Barcelona, Spain E-mail:
vispo@portos.bio.ub.es
COMMENT NCBI gi: 1279419
FEATURES
source Location/Qualifiers
1..427
/organism="Echinococcus granulosus"
exon 1..141

Mar 25 02:55

US-08-644-289-5.lge

4

/gene="Hbx2"
/note="homeobox"
intron 142..370
/gene="Hbx2"
exon 371..427
/gene="Hbx2"
/note="homeobox"
BASE COUNT 120 a 100 c 118 g 89 t
ORIGIN

Query Match 100.0%; Score 10; DB 84; Length 427;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 338 aggcattgcct 347
|||||
Cp 10 AGGCATGCCT 1

RESULT 3
LOCUS MEV41091 956 bp DNA BCT 02-MAY-1996
DEFINITION Unidentified marine eubacterium Hetp16 16S ribosomal RNA gene
sequence.
ACCESSION U41091
NID 91293542
KEYWORDS
SOURCE unidentified marine eubacterium.
ORGANISM unidentified marine eubacterium
Eubacteria.

REFERENCE 1 (bases 1 to 956)
AUTHORS Weidner,S., Arnold,W. and Puehler,A.
TITLE Diversity of uncultured microorganisms associated with the seagrass
JOURNAL Halophila stipulacea estimated by restriction fragment length
polymorphism analysis of PCR-amplified 16S rRNA genes
Appl. Environ. Microbiol. 62 (3), 766-771 (1996)
2 (bases 1 to 956)
AUTHORS Weidner,S.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-1995) Stefan Weidner, Fakultät fuer Biologie,
Lehrstuhl fuer Genetik, Universitaet Bielefeld, Universitaetstr.
25, Bielefeld, D-33615, Germany
COMMENT NCBI gi: 1293542
FEATURES
source Location/Qualifiers
1..956
/organism="unidentified marine eubacterium"
/note="associated with the seagrass Halophila stipulacea"
/clone="Hetp16"
/note="related to 16S ribosomal RNAs from bacteria of the
Hyphomonas group"
/product="16S ribosomal RNA"
BASE COUNT 254 a 208 c 294 g 200 t
ORIGIN

Query Match 100.0%; Score 10; DB 82; Length 956;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 730 aggcattgcct 739
|||||
Cp 10 AGGCATGCCT 1

RESULT 4

Mar 25 02:55

US-08-644-289-5.rge

5

LOCUS ME041088 956 bp DNA BCT 02-MAY-1996
DEFINITION Unidentified marine eubacterium Hstp13 16S ribosomal RNA gene
sequence.
ACCESSION U41088
NID q1293539
KEYWORDS
SOURCE unidentified marine eubacterium.
ORGANISM unidentified marine eubacterium
Eubacteria.
REFERENCE 1 (bases 1 to 956)
AUTHORS Weidner, S., Arnold, W. and Puehler, A.
TITLE Diversity of uncultured microorganisms associated with the seagrass
Halophila stipulacea estimated by restriction fragment length
polymorphism analysis of PCR-amplified 16S rRNA genes
JOURNAL Appl. Environ. Microbiol. 62 (3), 766-771 (1996)
AUTHORS Weidner, S.
TITLE Direct Submision
COMMENT Submitted (21-NOV-1995) Stefan Weidner, Fakultät fuer Biologie,
Lehrstuhl fuer Genetik, Universitaet Bielefeld, Universitaetsstr.
25, Bielefeld, D-33615, Germany
NCBI gi: 1293539
FEATURES
source location/Qualifiers
1..956
/organism="unidentified marine eubacterium"
/note="associated with the seagrass Halophila stipulacea"
/clone="Hstp13"
rRNA
/note="related to 16S ribosomal RNAs from bacteria of the
Hyphomonas group"
/product="16S ribosomal RNA"
BASE COUNT 252 a 209 c 296 g 199 t
ORIGIN
Query Match 100.0%; Score 10; DB 82; Length 956;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 730 agcagtcct 739
|||||
Cp 10 AGCATGCCT 1
RESULT 5
LOCUS HUMIGNT 2060 bp DNA PRI 10-APR-1996
DEFINITION Homo sapiens beta-1,6-N-acetylglucosaminyltransferase (Icgt) gene,
exon 1.
ACCESSION I41605
NID q9886291
KEYWORDS beta-1,6-N-acetylglucosaminyltransferase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2060)
AUTHORS Biehnhuizen, M.F., Maemura, K., Kudo, S. and Fukuda, M.
TITLE Genomic organization of core 2 and I branching
beta-1,6-N-acetylglucosaminyltransferases. Implication for
evolution of the beta-1,6-N-acetylglucosaminyltransferase gene
family
JOURNAL Glycobiology 5 (4), 417-425 (1995)
MEDLINE 96078409
COMMENT NCBI gi: 886291
FEATURES
Location/Qualifiers

Mar 25 02:55

US-08-644-289-5.rge

6

source 1..2060
/organism="Homo sapiens"
/tissue type="placenta"
312..1938
exon /gene="Icgt"
/EC number="3.2.1.23"
/note="Initiation codon starts at 1020; the I-branching
enzyme"
/number=1
5'UTR /codon_start=1
/product="beta-1,6-N-acetylglucosaminyltransferase"
312..1019
intron /gene="Icgt"
1939..>2060
/number=1
BASE COUNT 579 a 465 c 498 g 518 t
ORIGIN
Query Match 100.0%; Score 10; DB 90; Length 2060;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1016 agcagtcct 1025
|||||
Cp 10 AGCATGCCT 1
RESULT 6
LOCUS HSF3A120 2613 bp RNA PRI 22-APR-1996
DEFINITION H.sapiens mRNA for splicing factor SF3A120.
ACCESSION X85237
NID g989297
KEYWORDS SF3A120 gene; splicing factor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2613)
AUTHORS Kramer, A., Mulhauser, F., Wersig, C., Groning, K. and Bilbe, G.
TITLE Mammalian splicing factor SF3A120 represents a new member of the
SURP family of proteins and is homologous to the essential splicing
factor PRP21p of Saccharomyces cerevisiae
JOURNAL RNA 1 (3), 260-272 (1995)
MEDLINE 96079958
REFERENCE 2 (bases 1 to 2613)
AUTHORS Kramer, A.J.
TITLE Direct Submision
COMMENT Submitted (09-MAR-1995) A.J. Kramer, Universite de Geneve, Dept de
Biologie Cellulaire, 30 quai Ernest-Ansermet, 1211 Geneve 4,
SWITZERLAND
Sequence overlapping with the one under the acc# T75051.
NCBI gi: 899297
FEATURES
Location/Qualifiers
1..2613
/organism="Homo sapiens"
/cell_line="HeLa"
/clone_lib="lambda gt11"
/chromosome="22"
98..2479
/evidence=experimental
/gene="SF3A120"
/note="NCBI gi: 899298"
/codon_start=1

```
/product="human splicing factor"
/db_xref="PID:g899298"
/translation="MPAGPVQAVPPPPVPTEDTEEPASKEEDSAPSKPVVGIIT
PEPEVRNIVDKTASVANGCEFEARIRQNEINNPEFVLPNDPVIHAYVNHVSEFK
EKGAEQSAIPVWQ000QTQ00LPKQVQAVIQETTVKPEPPEEFIADEPSIS
AFDLVVKLTADQFVANGROFLTOLMOKEDRNVOFDELRPHSLFVFTKLVQYTKI
LIPKGLFSKLKKEAEAREVLDQVCYVMAKFOERERKKEEKEKERYANAQIDM
HOFVVEITVDQNEQGNPPPTPELGARILIQERYEFGSEEEVMEVESDEEDD
KOEKKEEPPSQLDQDTQVOMDEGSDDEEGQGNPPPEPMPPIPLPDPQVYAKD
YPRKASRPLPAPADEVISPITGEKIPASKMOEHMRTGLIDRMLEQDRBSIREKO
SDDEVYAPGLDISSLKQLAERTDIFVETETIIGKTIIGEELQREEEKVTMDSSGS
MAR7Q0AQA0NIT1AE0LEA1HKAKGLVPEDETKI1GPSKPNELIQOPEPPSANNI
PPSAPLTSVPRPPTPPVRTVWAVPVPWPPRPAWVRLPPGSLAPWPIIHAP
RINVPVWPPSAPPIMAPRPPPMVTPVAVVAPVAPVAPAPAPPPVHPPEMDEPTS
KKLKTEDSLMEPEERLARNKGPVSIKVQVPMQDKTEKMLNGQVLTFTPLTDQSVI
KYKIHENATGMPAGKQKIQYEGIFIKDSNSLAAYNANGAVTHALKERGRKK"
misc_feature
    /gene="SF3a120"
    /note="SURP module 1"
    251..379
misc_feature
    /gene="SF3a120"
    /note="SURP module 1"
    593..721
misc_feature
    /gene="SF3a120"
    /note="SURP module 2"
    857..901
misc_feature
    /gene="SF3a120"
    /note="charged domain"
    1708..1892
misc_feature
    /gene="SF3a120"
    /citation={2}
    2237..2467
misc_feature
    /gene="SF3a120"
    /note="ubiquitin-like domain"
    /note="ubiquitin-like domain"
BASE COUNT    659 a    800 c    679 g    475 t
ORIGIN

Query Match          100.0%; Score 10; DB 90; Length 2613;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      2341 aggcattgcct 2350
      |||||||||
Cp      10 AGGCATGCCT 1

RESULT 7
ID      574971      backbone; DNA; PBI; 3225 BP.
AC      574971;
DT      31-MAY-1995 (Rel. 43, Created)
DT      31-MAY-1995 (Rel. 43, Last updated, Version 1)
DE      STM=catecholamine-sulfating phenol sulfotransferase [human,
DE      Genomic, 3225 nt, segment 1 of 3].
KM
OS      Homo sapiens (human)
OC      Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC      Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
RN      (1)
RP      1-3225
RX      MEDLINE; 95100963.
RA      Dooley T.P., Probst P., Munroe P.B., Mole S.E., Liu Z.,
RA      Doggett N.A.;
RT      "Genomic organization and DNA sequence of the human
RT      catecholamine-sulfating phenol sulfotransferase gene (STM)",
RL      Biochem. Biophys. Res. Commun. 205:1325-1332(1994).
CC      NCBI gi: 833898
FH      Key      Location/Qualifiers
```

```
FH
FT      source      1..3225
FT
FT      /organism="Homo sapiens"
FT      /note="human"
FT      join(5..152,257..382,472..569,1964..2090,2186..2280,2779..
FT      CDS      2959,3082..3194)
FT      /gene="STM"
FT      /note="Description: catecholamine-sulfating phenol
FT      sulfotransferase"
FT      /codon_start=1
FT      /db_xref="PID:e163343"
FT      /translation="MEDIDPSTRPLEYKGVPLIKYFAEALGP1QSFQARDPDLIN
FT      YKSGCTTWVSQILDMITQGGDLEKCNRAPIYVRPDFLEVNDGEEPSGLETLKDPPEPR
FT      L
FT      IKSHLPALLPQTLLDQKVVVYVYARPKQVAVSYHFRMEKAHPERGTWDSFLEKF
FT      M      AGEVSYGSMYQHVQEMWELSRTHPVLYLFYEDMKENPREIQKILFVGRSLPEETMD
FT      F      MWQHTSPKEMKKNDMNTYTTVPQELMDHSISPFMRKGMADKTTFTYAQNERPDADY
FT      A      EKWAGCSLSPRSEL"
FT      SQ      Sequence 3225 BP; 677 A; 940 C; 905 G; 703 T; 0 other;

Query Match          100.0%; Score 10; DB 96; Length 3225;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      1708 aggcattgcct 1717
      |||||||||
Cp      10 AGGCATGCCT 1

RESULT 8
LOCUS      DROADHB039      3450 bp      DNA      INV      07-MAY-1996
DEFINITION      Drosophila pseudoobscura (subsp. bogotana, strain BOC 3389.9)
ACCESSION      M60996
KEYWORDS      g156855
SOURCE      alcohol dehydrogenase (adh) gene, complete cds.
ORGANISM      Drosophila pseudoobscura
      Eukaryota; Eukaryota; Eukaryota; Metazoa; Arthropoda;
      Insecta; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
      Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 3450)
AUTHORS      Schaeffer,S.W. and Miller,E.L.
TITLE      Nucleotide sequence analysis of Adh genes estimates the time of
      geographic isolation of the Bogota population of Drosophila
      pseudoobscura
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 88 (14), 6097-6101 (1991)
MEDLINE      91296768
COMMENT      NCBI gi: 156855
FEATURES      Location/Qualifiers
      exon      1..3450
      /organism="Drosophila pseudoobscura"
      /sub_species="bogotana"
      /tissue lib="EMBL4"
      <1..2206
      /note="unidentified reading frame"
      /number=1
      /codon_start=1
      8..14
      promoter
```

/gene="adh"
/note="adult"
40..119
/gene="adh"
/note="adult"
/codon_start=1
join(40..119,1062..1466,1527..1942)
/gene="adh"
/note="adult"
/product="alcohol dehydrogenase"
120..892
/gene="adh"
860..998
/gene="adh"
/note="larval"
/number=1
/codon_start=1
join(860..998,1062..1466,1527..1942)
/gene="adh"
/note="larval"
/product="alcohol dehydrogenase"
893..998
/gene="adh"
/note="adult"
/codon_start=1
join(893..998,1062..1466,1527..1942)
/gene="adh"
/note="adult"
/product="alcohol dehydrogenase"
join(906..998,1062..1466,1527..1793)
/gene="adh"
/EC_number="1.1.1.1"
/note="NCBI gi: 156856"
/codon_start=1
/product="alcohol dehydrogenase"
/db_xref="PID:q156856"
/translation="MSLTNNVVFVAGGIGDTSRELYKRNKIVYIDRIDNPA
IALKAINPVTITFPYDVTVPAETTKLKTFEAVKTIIDVILNAGGILDHQR
TIAVVTGIVNTTALIDFMDKRGPGGIIINISVTGFNATYQVYSGSRAAVN
ETSLAKIAPITCVTAATVNPQITKTTLVHKENSIDEPNVAKILEHPYQSQCA
ENFVAIEIKNNGAIIKLDGTELEPTTWQHMDSGI"
999..1061
/gene="adh"
/number=1
1062..1466
/gene="adh"
/number=2
/codon_start=1
1467..1526
/gene="adh"
/number=2
1527..1942
/gene="adh"
/number=3
/codon_start=1
1919..1924
/gene="adh"
/note="potential"
1942
/gene="adh"
2207..2507
/note="unidentified reading frame"
/number=1
2508..2912

/note="unidentified reading frame"
/number=2
/codon_start=1
2913..2970
/note="unidentified reading frame"
/number=2
2971..>3331
/note="unidentified reading frame"
/number=3
/codon_start=1
3326..3331
/note="unidentified reading frame; potential"
BASE COUNT 961 a 788 c 791 g 910 t
ORIGIN Chromosome IV, section 88.
Query Match 100.0%; Score 10; DB 83; Length 3450;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3306 aggcatagct 3315
|||||
Cp 10 AGGCATGCCCT 1
RESULT 9
LOCUS DROADHAI4 3463 bp DNA INV 07-MAY-1996
DEFINITION Drosophila pseudoobscura (strain AH144) alcohol dehydrogenase (adh)
gene, complete cds.
ACCESSION M60986
NID g156821
KEYWORDS alcohol dehydrogenase; alternative splicing.
SOURCE Drosophila pseudoobscura (tissue library: EMBL4) DNA.
ORGANISM Drosophila pseudoobscura
Eukaryota; Eukaryota; Eukaryota; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 3463)
Schaeffer,S.W. and Miller,E.L.
Nucleotide sequence analysis of Adh genes estimates the time of
geographic isolation of the Bogota population of Drosophila
pseudoobscura
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (14), 6097-6101 (1991)
MEDLINE 91296768
COMMENT NCBI gi: 156821
FEATURES
Location/Qualifiers
1..3463
/organism="Drosophila pseudoobscura"
/tissue_11b="EMBL4"
<1..2224
/note="unidentified reading frame"
/number=1
/codon_start=1
8..14
/gene="adh"
/note="adult"
40..119
/gene="adh"
/note="adult"
/codon_start=1
join(40..119,1082..1486,1547..1962)
/gene="adh"
/note="adult"
/product="alcohol dehydrogenase"
120..912

```
/gene="adh"  
880..1018  
/gene="adh"  
/note="larval"  
/number=1  
/codon_start=1  
join(880..1018,1082..1486,1547..1962)  
/gene="adh"  
/note="larval"  
/product="alcohol dehydrogenase"  
913..1018  
/gene="adh"  
/note="adult"  
/codon_start=1  
join(913..1018,1082..1486,1547..1962)  
/gene="adh"  
/note="adult"  
/product="alcohol dehydrogenase"  
join(926..1018,1082..1486,1547..1813)  
/gene="adh"  
/EC_number="1.1.1.1"  
/note="NCBI gi: 156822"  
/codon_start=1  
/product="alcohol dehydrogenase"  
/db_xref="PID:g156822"  
/translation="MSLTNRNVFVAGIGIGIDTSRELVRNIAKNLVILDIRIDNPAA  
IAEKAINPKVTITFYDYTVVPAVETTKILKTIENQKTIIDLINGAGILIDHQIER  
TIAVYTGIVNTTALIDFMDKRGKGGPCIIINI GSVGENALIVQPVYSSGAAYN  
FTSSIAKLAPITGVATYVNP GIKTTILVHKFSNIDVEPRVAEKLLEHPTQSQCA  
ENFVAIEIKNGAIAIKLIDIGTLEPIITWQHMSGI"  
1019..1081  
/gene="adh"  
/number=1  
1082..1486  
/gene="adh"  
/number=2  
/codon_start=1  
1487..1546  
/gene="adh"  
/number=2  
1547..1962  
/gene="adh"  
/number=3  
/codon_start=1  
1939..1944  
/gene="adh"  
/note="potential"  
1962  
/gene="adh"  
2225..2520  
/note="unidentified reading frame"  
/number=1  
2521..2925  
/note="unidentified reading frame"  
/number=2  
/codon_start=1  
2926..2983  
/note="unidentified reading frame"  
/number=2  
2984..3344  
/note="unidentified reading frame"  
/number=3  
/codon_start=1  
3339..3344  
polyA_signal
```

```
/note="potential; unidentified reading frame"  
BASE COUNT 967 a 792 c 795 g 909 t  
ORIGIN Chromosome IV, section 88.  
Query Match 100.0%; Score 10; DB 83; Length 3463;  
Best Local Similarity 100.0%; Pred. No. 8.12e+02;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 3319 aggcagtcct 3328  
Cp 10 AGGCATGCCT 1  
RESULT 10  
LOCUS DROADHAH15 3465 bp DNA INV 07-MAY-1996  
DEFINITION Drosophila pseudoobscura (strain AH135) alcohol dehydrogenase (adh)  
gene, complete cds.  
ACCESSION M60985  
NID g156823  
KEYWORDS alcohol dehydrogenase; alternative splicing.  
SOURCE Drosophila pseudoobscura (tissue library: EMBL4) DNA.  
ORGANISM Drosophila pseudoobscura  
Eukaryota; Eukaryota; Eukaryota; Metazoa; Arthropoda;  
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 3465)  
AUTHORS Schaeffer,S.W. and Miller,E.L.  
TITLE Nucleotide sequence analysis of Adh genes estimates the time of  
geographic isolation of the Bogota population of Drosophila  
pseudoobscura  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (14), 6097-6101 (1991)  
MEDLINE 91296768  
COMMENT NCBI gi: 156823  
FEATURES  
Location/Qualifiers  
1..3465  
1..3465  
/organism="Drosophila pseudoobscura"  
/tissue_lib="EMBL4"  
1..2227  
/note="unidentified reading frame"  
/number=1  
/codon_start=1  
8..14  
/gene="adh"  
/note="adult"  
40..119  
/gene="adh"  
/note="adult"  
/codon_start=1  
join(40..119,1084..1488,1551..1966)  
/gene="adh"  
/note="adult"  
/product="alcohol dehydrogenase"  
120..914  
/gene="adh"  
882..1020  
/gene="adh"  
/number=1  
/codon_start=1  
join(882..1020,1084..1488,1551..1966)  
/gene="adh"  
/note="larval"  
/product="alcohol dehydrogenase"  
915..1020  
/gene="adh"  
exon
```

```
/note="adult"
/codon_start=1
join(915..1020,1084..1488,1551..1966)
/gene="adh"
/note="adult"
/product="alcohol dehydrogenase"
join(928..1020,1084..1488,1551..1817)
/gene="adh"
/EC_number="1.1.1.1"
/note="NCBI gi: 156824"
/codon_start=1
/product="alcohol dehydrogenase"
/db_xref="PID:q156824"
/translation="MSLNKNVVFVAGLGIGDTSRELYKRNKLVYIDRIDPAA
IALEKAINPVTITFYPDYVVAETTKLAKTIFAOVKTIIDVLINGAILDHQIER
TIAVYTGIVTTTALIDFMDKRGPGCIIINISVTFNALIYQVYSGSKAAVN
FHSIAKLAPITGVATYVNPCTTKTTLVKEFNSMLDPERVAEKLLEHPQTSSQCA
ENFVAILELNKGAIMKLDIGLEPTITWQHWSGI"
1021..1083
/gene="adh"
/number=1
/number=1
1084..1488
/gene="adh"
/number=2
/codon_start=1
1489..1550
/gene="adh"
/number=2
1551..1966
/gene="adh"
/number=3
/codon_start=1
1943..1948
/gene="adh"
/note="potential"
1966
/gene="adh"
2228..2522
/note="unidentified reading frame"
2523..2927
/number=1
/note="unidentified reading frame"
/number=2
/codon_start=1
2928..2985
/note="unidentified reading frame"
/number=2
2986..3346
/note="unidentified reading frame"
/number=3
/codon_start=1
3341..3346
/note="unidentified reading frame; potential"
BASE COUNT 974 a 791 c 793 g 907 t
ORIGIN Chromosome IV, section 88.

Query Match 100.0%; Score 10; DB 83; Length 3465;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3321 aggcatacct 3330
|||||
Cp 10 AGGCATGCCT 1
```

```
RESULT 11
LOCUS DROADAH16 3466 bp DNA INV 07-MAY-1996
DEFINITION Drosophila pseudoobscura (strain AH162) alcohol dehydrogenase (adh)
gene, complete cds.
ACCESSION M60987
NID g156825
KEYWORDS alcohol dehydrogenase; alternative splicing.
SOURCE Drosophila pseudoobscura (tissue library: EMBL4) DNA.
ORGANISM Drosophila pseudoobscura
Eukaryota; Eukaryota; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 3466)
REFERENCE
AUTHORS Schaeffer,S.W. and Miller,E.L.
TITLE Nucleotide sequence analysis of Adh genes estimates the time of
geographic isolation of the Bogota population of Drosophila
pseudoobscura
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (14), 6097-6101 (1991)
MEDLINE 91296768
COMMENT NCBI gi: 156825
FEATURES
Location/Qualifiers
1..3466
/organism="Drosophila pseudoobscura"
/tissue_11b="EMBL4"
<1..2225
/note="unidentified reading frame"
/number=1
/codon_start=1
8..14
/gene="adh"
/note="adh"
40..119
/gene="adh"
/note="adh"
/codon_start=1
join(40..119,1082..1486,1547..1964)
/gene="adh"
/note="adh"
/product="alcohol dehydrogenase"
120..912
/gene="adh"
/note="adh"
880..1018
/gene="adh"
/note="laval"
/number=1
/codon_start=1
join(880..1018,1082..1486,1547..1964)
/gene="adh"
/note="laval"
/product="alcohol dehydrogenase"
913..1018
/gene="adh"
/note="adh"
/codon_start=1
join(913..1018,1082..1486,1547..1964)
/gene="adh"
/note="adh"
/product="alcohol dehydrogenase"
join(926..1018,1082..1486,1547..1813)
/gene="adh"
/EC_number="1.1.1.1"
/note="NCBI gi: 156826"
```


Mar 25 02:55

US-08-644-289-5.fge

15

```
/codon_start=1
/product="alcohol dehydrogenase"
/db_xref="PID:g156826"
/translation="MSLTKRNVVFVAGLGIGLDTSRSLYKRNLYIILDRIDNPA
IAELKAINPVVITTFYPDYVTVPAETTKLTKTIFNAVKTIDVLINGAGILDHQIER
TIANVYTGAVNTTALIDFMDKRGKGGGIIICNI GSVTFGNALYOVVYSGSKAAVAVN
FTSSIAKLAPITGVATYVNP GCTTKTIVHKRNSMLDVERVRAEKLEHP TQTSQQCA
ENFVAEILNKGAIAIKLIDGLTEPIITWQHMDSGI"
1019..1081
intron
/gene="adh"
/number=1
1082..1486.
/exon
/gene="adh"
/number=2
/codon_start=1
1487..1546
/intron
/gene="adh"
/number=2
1547..1964
/exon
/gene="adh"
/number=3
/codon_start=1
1950..1955
/polyA_signal
/gene="adh"
/note="potential"
1973
polyA_site
/gene="adh"
2226..2523
intron
/note="unidentified reading frame"
/number=1
2524..2928
/exon
/note="unidentified reading frame"
/number=2
2929..2986
/intron
/codon_start=1
2987..3347
/exon
/note="unidentified reading frame"
/number=3
3357..3362
polyA_signal
/codon_start=1
3357..3362
BASE COUNT 973 a 788 c 793 g 912 t
ORIGIN Chromosome IV, section 88.

Query Match 100.0%; Score 10; DB 83; Length 3466;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3322 aggcattgcct 3331
|||||
Cp 10 AGGCATGCCT 1
```

```
RESULT 12
LOCUS DROADNAH6 3467 bp DNA INV 07-MAY-1996
DEFINITION Drosophila pseudoobscura (strain AH165) alcohol dehydrogenase (adh)
gene, complete cds.
ACCESSION M60988
NID g156831
KEYWORDS alcohol dehydrogenase; alternative splicing.
SOURCE Drosophila pseudoobscura (tissue library: EMBL4) DNA.
ORGANISM Drosophila pseudoobscura
```

Mar 25 02:55

US-08-644-289-5.fge

16

```
REFERENCE
AUTHORS Schaeffer,S.W. and Miller,E.L.
TITLE Nucleotide sequence analysis of Adh genes estimates the time of
geographic isolation of the Bogota population of Drosophila
pseudoobscura
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (14), 6097-6101 (1991)
MEDLINE 91296768
COMMENT NCB1 gi: 156831
FEATURES
source Location/Qualifiers
1..3467
/organism="Drosophila pseudoobscura"
/tissue_1ib="EMBL4"
<1..2217
/note="UTR"
/number=1
join(40..119,1084..1488,1551..1966)
/gene="adh"
/note="adult"
/product="alcohol dehydrogenase"
40..119
/exon
/gene="adh"
/note="adult"
/codon_start=1
120..914
/gene="adh"
/note="adult"
882..1020
/exon
/gene="adh"
/note="larval"
/number=1
/codon_start=1
join(882..1020,1084..1488,1551..1966)
/gene="adh"
/note="larval"
/product="alcohol dehydrogenase"
915..1020
/exon
/gene="adh"
/note="adult"
/codon_start=1
join(915..1020,1084..1488,1551..1966)
/gene="adh"
/note="adult"
/product="alcohol dehydrogenase"
join(928..1020,1084..1488,1551..1817)
/gene="adh"
/EC number="1.1.1.1"
/note="NCBI gi: 156832"
/codon_start=1
/product="alcohol dehydrogenase"
/db_xref="PID:g156832"
/translation="MSLTKRNVVFVAGLGIGLDTSRSLYKRNLYIILDRIDNPA
IAELKAINPVVITTFYPDYVTVPAETTKLTKTIFNAVKTIDVLINGAGILDHQIER
TIANVYTGAVNTTALIDFMDKRGKGGGIIICNI GSVTFGNALYOVVYSGSKAAVAVN
FTSSIAKLAPITGVATYVNP GCTTKTIVHKRNSMLDVERVRAEKLEHP TQTSQQCA
ENFVAEILNKGAIAIKLIDGLTEPIITWQHMDSGI"
1021..1083
/exon
/gene="adh"
/number=1
1084..1488
/exon
/gene="adh"
/number=2
```

```
exon
intron
```

intron /codon_start=1
1489..1550
/gene="adh"
/number=2
1551..1966
/gene="adh"
/number=3
polyA_signal /codon_start=1
1943..1948
/gene="adh"
/note="potential"
1966
polyA_site /gene="adh"
2218..2532
/note="unidentified reading frame"
2533..2937
/number=1
exon /note="unidentified reading frame"
/number=2
/codon_start=1
2938..2995
/note="unidentified reading frame"
/number=2
2996..3356
/note="unidentified reading frame"
/number=3
polyA_signal /codon_start=1
3351..3356
/note="potential; unidentified reading frame"
BASE COUNT 980 a 793 c 786 g 908 t
ORIGIN Chromosome IV, section 88.
Query Match 100.0%; Score 10; DB 83; Length 3467;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3331 aggcatagcct 3340
|||||
Cp 10 AGGCATGCCT 1
RESULT 13
LOCUS DROADHB035 3471 bp DNA INV 07-MAY-1996
DEFINITION Drosophila pseudoobscura (subsp. bogotana, strain B06 3389.5)
alcohol dehydrogenase (adh) gene, complete cds.
ACCESSION M60994
MID g156851
KEYWORDS alcohol dehydrogenase; alternative splicing.
SOURCE Drosophila pseudoobscura (sub_species bogotana) (tissue library:
EMBL4) DNA.
ORGANISM Drosophila pseudoobscura
Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 3471)
AUTHORS Schaeffer,S.W. and Miller,E.L.
TITLE Nucleotide sequence analysis of Adh genes estimates the time of
geographic isolation of the Bogota population of Drosophila
pseudobscura
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (14), 6097-6101 (1991)
MEDLINE 91296768
COMMENT NCBI gi: 156851
FEATURES Location/Qualifiers
source 1..3471

exon /organism="Drosophila pseudoobscura"
/sub_species="Bogotana"
/tissue_lbb="EMBL4"
<1..2232
/note="unidentified reading frame"
/number=1
/codon_start=1
8..14
/gene="adh"
/note="adult"
40..119
/gene="adh"
/note="adult"
/codon_start=1
join(40..119,1083..1487,1550..1968)
/gene="adh"
/note="adult"
/product="alcohol dehydrogenase"
120..914
/gene="adh"
/note="adult"
882..1020
/gene="adh"
/note="larval"
/number=1
/codon_start=1
join(882..1020,1083..1487,1550..1968)
/gene="adh"
/note="larval"
/product="alcohol dehydrogenase"
915..1020
/gene="adh"
/note="adult"
/codon_start=1
join(915..1020,1083..1487,1550..1968)
/gene="adh"
/note="adult"
/product="alcohol dehydrogenase"
join(928..1020,1083..1487,1550..1816)
/gene="adh"
/EC_number="1.1.1.1"
/note="NCBI gi: 156852"
/product="alcohol dehydrogenase"
/db_xref="PID:g156852"
/translation="MSLTNKNVFPVAGIGIGIDTSRELVKRNLTKNLVIIDRIIDPAA
IAELKAINRVITTFPPYDVTPVAETTKLKTTPAQVKTIIVLINSAGILDDHDIET
TIANVYGLVNTTAAIIDFMDRKGGPGGIIICNIGVGFNAIYQVPVYSGSKAAVN
FTTSIALPITGVTAATVPGITTITLVHKNSMLVEPRVAEKLEHPYTSQQA
ENFKAELKNGAIKMKLDGLTEP ITWYQHDSGI"
1021..1082
/gene="adh"
/number=1
1083..1487
/gene="adh"
/number=2
/codon_start=1
1488..1549
/gene="adh"
/number=2
1550..1968
/gene="adh"
/number=3
/codon_start=1

polyA_signal 1945..1950
/gene="adh"
/note="potential"
1968
polyA_site
/gene="adh"
2233..2528
intron
/note="unidentified reading frame"
/number=1
2529..2933
exon
/note="unidentified reading frame"
/number=2
/codon_start=1
2934..2991
intron
/note="unidentified reading frame"
/number=2
2992..3352
exon
/note="unidentified reading frame"
/number=3
/codon_start=1
polyA_signal 3347..3352
/note="potential; unidentified reading frame"
BASE COUNT 981 a 791 c 793 g 906 t
ORIGIN Chromosome IV, section 88.
Query Match 100.0%; Score 10; DB 83; Length 3471;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3327 aggcattcct 3336
|||||
Cp 10 AGGCATGCTT 1
RESULT 14
LOCUS DROADH031 3472 bp DNA INV 07-MAY-1996
DEFINITION Drosophila pseudoobscura (subsp. bogotana, strain BOC 3389.1)
alcohol dehydrogenase (adh) gene, complete cds.
ACCESSION M60990
NID g156843
KEYWORDS alcohol dehydrogenase; alternative splicing.
SOURCE Drosophila pseudoobscura (sub_species bogotana) (tissue library:
EMBL4) DNA.
ORGANISM Drosophila pseudoobscura
Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Diptera; Brachyera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 3472)
REFERENCE
AUTHORS Schaeffer,S.W. and Miller,E.L.
TITLE Nucleotide sequence analysis of Adh genes estimates the time of
geographic isolation of the Bogota population of Drosophila
pseudoobscura
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (14), 6097-6101 (1991)
MEDLINE 91296768
COMMENT NCBI gi: 156843
FEATURES
Location/Qualifiers
1..3472
/organism="Drosophila pseudoobscura"
/sub_species="bogotana"
/tissue_lib="EMBL4"
<1..2229
exon
/note="unidentified reading frame"
/number=1
/codon_start=1
8..14
promoter

exon
/gene="adh"
/note="adult"
40..119
/gene="adh"
/note="adult"
/codon_start=1
join(40..119,1085..1489,1550..1965)
mRNA
/gene="adh"
/note="adult"
/product="alcohol dehydrogenase"
120..915
intron
/gene="adh"
883..1021
exon
/gene="adh"
/note="larval"
/number=1
/codon_start=1
join(883..1021,1085..1489,1550..1965)
mRNA
/gene="adh"
/note="larval"
/product="larval"
/product="alcohol dehydrogenase"
916..1021
exon
/gene="adh"
/note="adult"
/codon_start=1
join(916..1021,1085..1489,1550..1965)
mRNA
/gene="adh"
/note="adult"
/product="alcohol dehydrogenase"
join(929..1021,1085..1489,1550..1816)
/gene="adh"
/EC_number="1.1.1.1"
/note="NCBI gi: 156844"
/codon_start=1
/product="alcohol dehydrogenase"
/db_xref="PID:g156844"
/translation="MSLTKNVVFVAGLGICGLDTSRELYKRNILKNVILDIRIDNPA
IAELKAINPKVITFYPYDVTPVAETPKLKTTFQAVTIDVLINGAGILDDHQIER
TIAVNYGLNNTTALIDFMDRKRGGGGLINIGSVTFNATYOVVYSGSKAAVN
FTSSIAKLAPITGVATYTPPCITKTVLHKNSHIDVEPVAEKLELPYOTSQQA
ENFVKALELNKGAIMKLDLGTLEPTWQHWSGI"
1022..1084
intron
/gene="adh"
/number=1
1085..1489
exon
/gene="adh"
/number=2
/codon_start=1
1490..1549
intron
/gene="adh"
/number=2
1550..1965
exon
/gene="adh"
/number=3
/codon_start=1
1942..1947
polyA_signal
/gene="adh"
/note="potential"
1965
polyA_site
/gene="adh"
2230..2529
intron
/note="unidentified reading frame"
/number=1
2530..2934
exon

```
/note="unidentified reading frame"
/number=2
/codon start=1
intron      2935..2992
            /note="unidentified reading frame"
            /number=2
            /note="unidentified reading frame"
            /number=3
            /codon start=1
            polyA_signal      3348..3353
                                /note="unidentified reading frame; potential"
BASE COUNT      972 a      793 c      794 g      913 t
ORIGIN      Chromosome IV, section 88.

Query Match      100.0%; Score 10; DB 83; Length 3472;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      3328 aggcatagcct 3337
        |||||||||
Cp      10 AGGCATGCCT 1

RESULT 15
LOCUS      DROHMAH10 3472 bp DNA INV 07-MAY-1996
DEFINITION Drosophila pseudoobscura (strain MH100) alcohol dehydrogenase adh
            gene, complete cds.
ACCESSION      M60982
NID      9156815
KEYWORDS      alcohol dehydrogenase; alternative splicing.
SOURCE      Drosophila pseudoobscura (tissue library: EMBL4) DNA.
ORGANISM      Drosophila pseudoobscura
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;
            Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 3472)
            Schaeffer,S.W. and Miller,E.L.
            Nucleotide sequence analysis of Adh genes estimates the time of
            geographic isolation of the Bogota population of Drosophila
            pseudoobscura
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 88 (14), 6097-6101 (1991)
MEDLINE      91296768
COMMENT      NCBI gi: 156815
FEATURES
            Location/Qualifiers
            1..3472
            /organism="Drosophila pseudoobscura"
            /tissue_lib="EMBL4"
            <1..2228
            /note="unidentified reading frame"
            /number=1
            /codon start=1
            8..14
            /gene="adh"
            /note="adh"
            /note="adult"
            join(40..119,1084..1488,1549..1964)
            /gene="adh"
            /note="adult"
            /product="alcohol dehydrogenase"
            40..119
            /gene="adh"
            /note="adh"
            /note="adult"
            /codon start=1
            120..914
            intron
```

```
/gene="adh"
/note="adult"
exon      882..1020
            /gene="adh"
            /note="larval"
            /number=1
            /codon start=1
            join(882..1020,1084..1488,1549..1964)
            /gene="adh"
            /note="larval"
            /product="alcohol dehydrogenase"
            join(915..1020,1084..1488,1549..1964)
            /gene="adh"
            /note="adult"
            /product="alcohol dehydrogenase"
            915..1020
            /gene="adh"
            /note="adult"
            /codon start=1
            join(928..1020,M60981:1084..1488,1549..1815)
            /EC_number="1.1.1.1"
            /note="NCBI gi: 1303659"
            /codon start=1
            /product="alcohol dehydrogenase"
            /db_xref="PID:g1303659"
            /translation="MSLTKNNVVFVAGLGIGLDTSRBLKGNLKNLVLIDRLIDNPAA
            IAEKKAINKRYTITFFPYDYTPVPAETTKLKTIFAQYKTTIVLINAAGLDDHIER
            TIAVNYGLVNTTATILDFMDRRKGGPGGICNIGSVTFGNALIQVPVYSGKAAYN
            FTSSIAKLAPLTGVTATVNPGLTKTLVHKNSMLDVEPRVAEKLEHPYTSQQCA
            ENFYKALEIKRGAIMKLDLGTIEPTITVQHMDSGI"
            1021..1083
            /gene="adh"
            /number=1
            1084..1488
            /gene="adh"
            /number=2
            /codon start=1
            1489..1548
            /gene="adh"
            /number=2
            1549..1964
            /gene="adh"
            /number=3
            /codon start=1
            1941..1946
            /gene="adh"
            /note="potential"
            1964
            /gene="adh"
            2229..2529
            /note="unidentified reading frame"
            /number=1
            2530..2934
            /note="unidentified reading frame"
            /number=2
            /codon start=1
            2935..2992
            /note="unidentified reading frame"
            /number=2
            2993..>3353
            /note="unidentified reading frame"
            /number=3
            /codon start=1
            3348..3353
            polyA_signal
```

Mar 25 02:55

US-08-644-289-5.rge

23

BASE COUNT 967 a 793 c 800 g 912 t /note="potential; unidentified reading frame"
ORIGIN Chromosome IV, section 88.

Query Match 100.0%; Score 10; DB 83; Length 3472;
Best Local Similarity 100.0%; Pred. No. 8.12e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3328 aggcattgcct 3337
 |||||
Cp 10 AGGCATTGCT 1

Search completed: Tue Mar 25 02:57:03 1997
Job time : 55 secs.

Mar 25 02:55

US-08-644-289-5.mrg

3

CC sequences are found near the origin of replication of certain animal
CC viruses and animal cells. Four mutant forms of p53 protein which are
CC commonly found in human tumours do not have the ability to bind to
CC these sequences. Thus a function of p53 may be mediated by its
CC ability to bind specific DNA sequences in the human genome. The
CC sequence shown is a consensus sequence for p53 DNA binding. When
CC inserted upstream and adjacent to a reporter gene the sequence
CC allows identification of wild type p53, and such a construct could
CC be used for diagnosis of p53 mutations and onset and development of
CC various cancers. The construct may also be used to screen potential
CC chemotherapeutic agents and to identify agents which specifically
CC bind p53-specific DNA. Also wild-type p53 gene function may be
CC restored to neoplastic cells having a mutation in their p53 gene.
SQ See also Q31949-84.
Sequence 10 BP; 0 A; 1 C; 1 G; 0 U;

Query Match 100.0%; Score 10; DB 5; Length 10;
Best Local Similarity 20.0%; Pred. No. 1.99e+02;
Matches 2; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

Db 1 tttcwwggyy 10
:::|::|::|
Qy 1 AGCGATGCCT 10

RESULT 2
ID Q31948 standard; DNA; 10 BP.
AC Q31948;
DT 27-APR-1993 (first entry)
DE Monomeric p53-specific DNA binding site.
KM P53; DNA-binding; cancer; neoplasia; tumour; ds.
OS Synthetic.
PN EP-518650-A.
PD 16-DEC-1992.
PF 10-JUN-1992; 305333.
PR 14-JUN-1991; US-715182.
PR 31-MAR-1992; US-860758.
PA (PHAR-) PHARMAGENICS INC.
PA (UYX0) UNIV JOHNS HOPKINS.
PI Kinzler KM, Sherman MI, Vogelstein B;
DR WPI; 92-417505/51.
PT Detection and expression of wild type P53 protein - useful for
PT diagnosing and treating cancers, and for screening potential
PT chemotherapeutic agents.
PS Claim 22, Page 24; 51pp; English.
CC Wild-type p53 protein binds specific fragments of human chromosomal
CC DNA. Each fragment contains no more than one monomer of the double
CC stranded sequence shown separated by 0-13 base pairs. Some of these
CC sequences are found near the origin of replication of certain animal
CC viruses and animal cells. Four mutant forms of p53 protein which are
CC commonly found in human tumours do not have the ability to bind to
CC these sequences. Thus a function of p53 may be mediated by its
CC ability to bind specific DNA sequences in the human genome. The
CC sequence shown is a consensus sequence for p53 DNA binding. When
CC inserted upstream and adjacent to a reporter gene the sequence
CC allows identification of wild type p53, and such a construct could
CC be used for diagnosis of p53 mutations and onset and development of
CC various cancers. The construct may also be used to screen potential
CC chemotherapeutic agents and to identify agents which specifically
CC bind p53-specific DNA. Also wild-type p53 gene function may be
CC restored to neoplastic cells having a mutation in their p53 gene.
SQ See also Q31949-84.
Sequence 10 BP; 0 A; 1 C; 1 G; 0 U;

Mar 25 02:55

US-08-644-289-5.mrg

4

Query Match 100.0%; Score 10; DB 5; Length 10;
Best Local Similarity 20.0%; Pred. No. 1.99e+02;
Matches 2; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

Db 1 tttcwwggyy 10
:::|::|::|
Cp 10 AGCGATGCCT 1

RESULT 3
ID T10102 standard; DNA; 20 BP.
AC T10102;
DT 17-JUL-1996 (first entry)
DE Human cholecystokinin-B/gastrin receptor PCR primer.
KM Cholecystokinin B; CCK-B; gastrin; receptor; human; isoform;
KM antagonist; gastric acid secretion; inhibition; anti-anxiety; ss.
OS Synthetic.
PN J07324099-A.
PD 12-DEC-1995.
PF 30-MAY-1994; 117015.
PR 30-MAY-1994; JP-117015.
PA (YAMA) YAMANOUCHI PHARM CO LTD.
DR WPI; 96-065478/07.

PT Cholecystokinin-B/gastrin receptor protein and gene - used to screen
PT and evaluate agents for use as gastric acid secretion inhibitors and
PT anti-anxiety agents
PS Example 2; Page 7; 14pp; Japanese.
CC A cDNA sequence encoding the human isoform of cholecystokinin-B/
CC gastrin receptor was isolated (see T10094). The present sequence is
CC that of a PCR primer which was used in cloning the full-length cDNA.
CC Cells transfected with the human coding sequence will express the
CC receptor on their surface and will thus be useful for screening
CC candidate antagonists of the receptor. Such antagonists have
CC potential use as inhibitors of gastric acid secretion or as anti-
CC anxiety agents.
SQ Sequence 20 BP; 4 A; 6 C; 4 G; 6 T;

Query Match 100.0%; Score 10; DB 19; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.99e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 agcgatgcct 16
|||||||
Qy 1 AGCGATGCCT 10

RESULT 4
ID T15968 standard; cDNA to mRNA; 20 BP.
AC T15968;
DT 09-SEP-1996 (first entry)
DE Primer for human CCK-B/gastrin receptor gene amplification.
KM CCK-B/gastrin; cholecystokinin-B/gastrin receptor; antagonist;
KM gastric secretion; tranquilizer; assay; production; primer; PCR;
KM polymerase chain reaction; amplification; ss.
OS Synthetic.
PN J08027190-A.
PD 30-JAN-1996.
PF 22-JUL-1994; 170663.
PR 22-JUL-1994; JP-170663.
PA (YAMA) YAMANOUCHI PHARM CO LTD.
DR WPI; 96-136331/14.
PT Gene encoding cholecystokinin-B/gastrin receptor - useful, e.g. to
PT develop antagonists used to evaluate inhibitors of gastric acid
PT secretion

Mar 25 02:55

US-08-644-289-5.rmg

5

PS Example 2; Page 7; 15pp; Japanese.
CC T15963-68 were used to PCR amplify a DNA fragment contg. a human
CC cholecystokinin-B/gastrin receptor gene (see T15960). Host cells
CC transformed with the gene are useful for recombinant production of the
CC receptor protein. The receptor is useful for the development of
CC pharmaceutical drugs, esp. antagonists to evaluate inhibitors of gastric
CC acid secretion and tranquilizers, etc..

SQ Sequence 20 BP; 4 A; 6 C; 4 G; 6 T;

Query Match 100.0%; Score 10; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.99e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 aggcattgcct 16

0y 1 AGGCATTCCT 10

RESULT 5

ID 031978 standard; DNA; 20 BP.

AC 031978;

DT 27-APR-1993 (first entry)

DE Synthetic oligomer 4, to determine motif for binding to p53.

KM p53; DNA-binding; cancer; neoplasia; tumour; concatemer; ss.

OS Synthetic.

PN EP-518650-A.

PD 16-DEC-1992.

PF 10-JUN-1992; 305333.

PR 14-JUN-1991; US-715182.

PA 31-MAR-1992; US-860758.

PA (PHAR-) PHARMACENICS INC.

PA (UYJO) UNIV JOHNS HOPKINS.

PI Kinzler KM, Sherman MI, Vogelstein B;

DR WPI; 92-417505/51.

PT Detection and expression of wild type p53 protein - useful for
PT diagnosing and treating cancers, and for screening potential

PT chemotherapeutic agents

PS Example 10; Fig 10; 51pp; English.

CC Wild-type p53 protein binds specific fragments of human chromosomal

CC DNA. 23 clones which bound p53 were obtd. by whole genome PCR.

CC The sequences of these 23 clones was compared and an average insert

CC of 307 bp (range 139-470) was found. Only 18 of the clones contained

CC unique fragments. Alignment of the inserts revealed a striking and

CC consistent feature, each binding site contained two copies of the

CC 10 bp motif 5'-RRRCWGGY-3' separated by 0-13 bp. The 10 bp

CC consensus sequence contains internal symmetry. Synthetic oligomers

CC were synthesised to determine portions of the motif vital for p53

CC binding. Mutation analysis revealed that the dimer of the consensus

CC motif is required to bind p53, that G residues at nucleotides 4

CC and 6 are critical for p53 binding and that mutant p53 proteins

CC found in tumours do not bind to the consensus sequence.

CC Oligonucleotide 4 was capable of binding p53.

CC See also 031948-84.

SQ Sequence 20 BP; 4 A; 6 C; 6 G; 4 T;

Query Match 100.0%; Score 10; DB 5; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.99e+02;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 aggcattgcct 10

0y 1 AGGCATTCCT 10

Mar 25 02:55

US-08-644-289-5.rmg

6

RESULT 6

ID T15968 standard; cDNA to mRNA; 20 BP.

AC T15968;

DT 09-SEP-1996 (first entry)

DE Primer for human CCK-B/gastrin receptor gene amplification.

KM CCK-B/gastrin; cholecystokinin-B/gastrin receptor; antagonist;

KM gastric secretion; tranquilizer; assay; production; primer; PCR;

KM polymerase chain reaction; amplification; ss.

OS Synthetic.

PN J08027190-A.

PD 30-JAN-1996.

PF 22-JUL-1994; 170663.

PR 22-JUL-1994; JP-170663.

PA (YAMA) YAMANOUCHI PHARM CO LTD.

DR WPI; 96-136331/14.

PT Gene encoding cholecystokinin-B/gastrin receptor - useful, e.g. to

PT develop antagonists used to evaluate inhibitors of gastric acid

PT secretion

PS Example 2; Page 7; 15pp; Japanese.

CC T15963-68 were used to PCR amplify a DNA fragment contg. a human

CC cholecystokinin-B/gastrin receptor gene (see T15960). Host cells

CC transformed with the gene are useful for recombinant production of the

CC receptor protein. The receptor is useful for the development of

CC pharmaceutical drugs, esp. antagonists to evaluate inhibitors of gastric

CC acid secretion and tranquilizers, etc..

SQ Sequence 20 BP; 4 A; 6 C; 4 G; 6 T;

Query Match 100.0%; Score 10; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.99e+02;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 aggcattgcct 16

Cp 10 AGGCATTCCT 1

RESULT 7

ID 031978 standard; DNA; 20 BP.

AC 031978;

DT 27-APR-1993 (first entry)

DE Synthetic oligomer 4, to determine motif for binding to p53.

KM p53; DNA-binding; cancer; neoplasia; tumour; concatemer; ss.

OS Synthetic.

PN EP-518650-A.

PD 16-DEC-1992.

PF 10-JUN-1992; 305333.

PR 14-JUN-1991; US-715182.

PA 31-MAR-1992; US-860758.

PA (PHAR-) PHARMACENICS INC.

PA (UYJO) UNIV JOHNS HOPKINS.

PI Kinzler KM, Sherman MI, Vogelstein B;

DR WPI; 92-417505/51.

PT Detection and expression of wild type p53 protein - useful for

PT diagnosing and treating cancers, and for screening potential

PT chemotherapeutic agents

PS Example 10; Fig 10; 51pp; English.

CC Wild-type p53 protein binds specific fragments of human chromosomal

CC DNA. 23 clones which bound p53 were obtd. by whole genome PCR.

CC The sequences of these 23 clones was compared and an average insert

CC of 307 bp (range 139-470) was found. Only 18 of the clones contained

CC unique fragments. Alignment of the inserts revealed a striking and

CC consistent feature, each binding site contained two copies of the

CC 10 bp motif 5'-RRRCWGGY-3' separated by 0-13 bp. The 10 bp

CC consensus sequence contains internal symmetry. Synthetic oligomers

CC were synthesised to determine portions of the motif vital for p53 binding. Mutation analysis revealed that the dimer of the consensus motif is required to bind p53, that G residues at nucleotides 4 and 6 are critical for p53 binding and that mutant p53 proteins CC found in tumours do not bind to the consensus sequence.
CC Oligonucleotide 4 was capable of binding p53.
CC See also Q31948-84.

Sequence 20 BP; 4 A; 6 C; 6 G; 4 T;

Query Match 100.0%; Score 10; DB 5; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.99e+02;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 agcgcgcct 10

|||||

Cp 10 AGCGATGCCT 1

RESULT 8

ID T110102 standard; DNA; 20 BP.

AC T110102;

DT 17-JUL-1996 (first entry)

DE Human cholecystokinin-B/gastrin receptor PCR primer.

KM Cholecystokinin B; CCK-B; gastrin; receptor; human; isoform;

KM antagonist; gastric acid secretion; inhibition; anti-anxiety; ss.

OS Synthetic.

PN J07324099-A.

PD 12-DEC-1995.

PF 30-MAY-1994; 117015.

PR 30-MAY-1994; JP-117015.

RA (YAMA) YAMANOUCHI PHARM CO LTD.

DR WPI; 96-065478/07.

PT Cholecystokinin-B/gastrin receptor protein and gene - used to screen

PT and evaluate agents for use as gastric acid secretion inhibitors and

PT anti-anxiety agents

PS Example 2; Page 7; 14pp; Japanese.

CC A cDNA sequence encoding the human isoform of cholecystokinin-B/

CC gastrin receptor was isolated (see T110094). The present sequence is

CC that of a PCR primer which was used in cloning the full-length cDNA.

CC Cells transformed with the human coding sequence will express the

CC receptor on their surface and will thus be useful for screening

CC candidate antagonists of the receptor. Such antagonists have

CC potential use as inhibitors of gastric acid secretion or as anti-

CC anxiety agents.

Sequence 20 BP; 4 A; 6 C; 4 G; 6 T;

Query Match 100.0%; Score 10; DB 19; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.99e+02;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 agcgcgcct 16

|||||

Cp 10 AGCGATGCCT 1

RESULT 9

ID T11108 standard; cDNA; 24 BP.

AC T11108;

DT 09-AUG-1996 (first entry)

DE Transforming growth factor beta cDNA clone MP-121 PCR primer L02.

KM TGF-beta; MP-121; mitogen; differentiation; induction; promotion;

KM maintenance; morphogen; tissue regeneration; dental implantation;

KM wound healing; polymerase chain reaction; PCR primer; ss.

OS Synthetic.

PN DE19511243-A1.

PD 04-JAN-1996.

PF 27-MAR-1995; 011243.

PR 01-JUL-1994; DE-423190.

PA (BIOP-) BIOPHARM GES BIOTECHNOLOGISCHEN ENTWICKL.

PI Bechtold R, Neidhardt H, Pohl J, Hoelten G;

DR WPI; 96-050788/06.

PT DNA encoding transforming growth factor beta MP-121 - has mitogenic

PT and differentiation-inducing activity, e.g. for use in wound healing

PS Example 1; Page 6; 15pp; German.

CC A cDNA library prepared using total RNA from human liver was

CC subjected to PCR amplification using primers corresp. to conserved

CC regions within the TGF-beta family. Amplification products were

CC subcloned and sequenced; one clone (designated pSK-MP121) was found

CC to contain a new sequence. Part of the insert from this clone was

CC used to re-screen the human liver cDNA library and a 2272 bp fragment

CC coding for a TGF-beta-like protein was isolated. The protein encoded

CC by the cDNA insert has mitogenic and differentiation-inducing

CC properties making it (or fusion proteins comprising it or heterodimers

CC of the protein with a cysteine knot motif protein) useful for inducing

CC tissue regeneration. The present sequence is that of a PCR primer

CC based on a Ddel fragment comprising nucleotides 931-1304 of the

CC 2272 bp insert. The Ddel fragment was used as a probe to screen the

CC human liver library; positive clones were then subjected to PCR

CC amplification using primer L02 (i.e. this sequence) and primer L011

(see T11109).

Sequence 24 BP; 6 A; 5 C; 7 G; 6 T;

Query Match 100.0%; Score 10; DB 20; Length 24;

Best Local Similarity 100.0%; Pred. No. 1.99e+02;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 agcgcgcct 17

|||||

Qy 1 AGCGATGCCT 10

RESULT 10

ID T11107 standard; cDNA; 24 BP.

AC T11107;

DT 09-AUG-1996 (first entry)

DE Transforming growth factor beta cDNA clone MP-121 internal primer.

KM TGF-beta; MP-121; mitogen; differentiation; induction; promotion;

KM maintenance; morphogen; tissue regeneration; dental implantation;

KM wound healing; polymerase chain reaction; PCR primer; ss.

OS Synthetic.

PN DE19511243-A1.

PD 04-JAN-1996.

PF 27-MAR-1995; 011243.

PR 01-JUL-1994; DE-423190.

PA (BIOP-) BIOPHARM GES BIOTECHNOLOGISCHEN ENTWICKL.

PI Bechtold R, Neidhardt H, Pohl J, Hoelten G;

DR WPI; 96-050788/06.

PT DNA encoding transforming growth factor beta MP-121 - has mitogenic

PT and differentiation-inducing activity, e.g. for use in wound healing

PS Example 1; Page 6; 15pp; German.

CC A cDNA library prepared using total RNA from human liver was

CC subjected to PCR amplification using primers corresp. to conserved

CC regions within the TGF-beta family. Amplification products were

CC subcloned and sequenced; one clone (designated pSK-MP121) was found

CC to contain a new sequence. Part of the insert from this clone was

CC used to re-screen the human liver cDNA library and a 2272 bp fragment

CC coding for a TGF-beta-like protein was isolated. The protein encoded

CC by the cDNA insert has mitogenic and differentiation-inducing

CC properties making it (or fusion proteins comprising it or heterodimers

CC of the protein with a cystine knot motif protein) useful for inducing
 CC tissue regeneration. The present sequence is that of an internal
 CC primer which was used for isolating the cDNA coding sequence.
 SQ Sequence 24 BP; 6 A; 5 C; 7 G; 6 T;

Query Match 100.0%; Score 10; DB 20; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.99e+02;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 agcagtcgct 17
 |||||
 Qy 1 AGCGATCGCT 10

RESULT 11
 ID 047714 standard; DNA; 24 BP.

AC 047714;

DT 11-FEB-1994 (first entry)

DE TGF-beta-like clone MP-121 internal nested primer.

KM Human; transforming growth factor; beta; TGF-beta; pharmaceutical;

KM bone; cartilage; tooth; wound repair; immunosuppressor; PCR; amplity;

KM organ transplant; cosmetic surgery; antibody; diagnosis; primer;

KM polymerase chain reaction; ss.

OS Homo sapiens.

PN W09316099-A.

PD 19-AUG-1993.

PF 12-FEB-1993; E00350.

PR 12-FEB-1992; EP-102324.

PA (BIOP-) BIOPHARM GES BIOTECHNOLOGISCHEN ENTWICKL.

PI Hoelten G, Neidhardt H;

DR WPI; 93-272824/34.

PT New transforming growth factor-beta family proteins and DNA -

PT used in tissue and wound repair, in treatment of bone, cartilage

PT and tooth defects, and antibodies for diagnosis

PS Example 1; Page 13; 29pp; English.

CC The sequences given in 047711-14 are primers which may be used in

CC the amplification of the liver derived human transforming growth

CC factor-beta (TGF-beta) clone MP-121. The amplification product

CC encodes a protein which may be used in a pharmaceutical composition

CC for the treatment of various bone, cartilage or tooth defects and in

CC tissue and wound repair processes. These proteins may also be used

CC as immunosuppressors in organ transplants and in cosmetic surgery.

CC Antibodies raised against these proteins may be used for diagnostic

CC purposes.

SQ Sequence 24 BP; 6 A; 5 C; 7 G; 6 T;

Query Match 100.0%; Score 10; DB 8; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.99e+02;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 agcagtcgct 17
 |||||
 Qy 1 AGCGATCGCT 10

RESULT 12
 ID T11108 standard; cDNA; 24 BP.

AC T11108;

DT 09-AUG-1996 (first entry)

DE Transforming growth factor beta cDNA clone MP-121 PCR primer 102.

KM TGF-beta; MP-121; mitogen; differentiation; induction; promotion;

KM maintenance; morphogen; tissue regeneration; dental implantation;

KM wound healing; polymerase chain reaction; PCR primer; ss.

OS Synthetic.

PN DE19511243-A1.

PD 04-JAN-1996.

PF 27-MAR-1995; 011243.

PR 01-JUL-1994; DE-423190.

PA (BIOP-) BIOPHARM GES BIOTECHNOLOGISCHEN ENTWICKL.

PI Bechtold R, Neidhardt H, Pohl J, Hoelten G;

DR WPI; 96-050788/06.

PT DNA encoding transforming growth factor beta MP-121 - has mitogenic

PT and differentiation-inducing activity, e.g. for use in wound healing

PS Example 1; Page 6; 15pp; German.

CC A cDNA library prepared using total RNA from human liver was

CC subjected to PCR amplification using primers corresp. to conserved

CC regions within the TGF-beta family. Amplification products were

CC subcloned and sequenced; one clone (designated pSK-MP121) was found

CC to contain a new sequence. Part of the insert from this clone was

CC used to re-screen the human liver cDNA library and a 2272 bp fragment

CC coding for a TGF-beta-like protein was isolated. The protein encoded

CC by the cDNA insert has mitogenic and differentiation-inducing

CC properties making it (or fusion proteins comprising it or heterodimers

CC of the protein with a cystine knot motif protein) useful for inducing

CC tissue regeneration. The present sequence is that of a PCR primer

CC based on a Ddel fragment comprising nucleotides 931-1304 of the

CC 2272 bp insert. The Ddel fragment was used as a probe to screen the

CC human liver library; positive clones were then subjected to PCR

CC amplification using primer 102 (i.e. this sequence) and primer 1011

CC (see T11109).

SQ Sequence 24 BP; 6 A; 5 C; 7 G; 6 T;

Query Match 100.0%; Score 10; DB 20; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.99e+02;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 agcagtcgct 17
 |||||
 Cp 10 AGCGATCGCT 1

RESULT 13

ID 080759 standard; DNA; 42 BP.

AC 080759;

DT 19-JUL-1995 (first entry)

DE Pilot oligonucleotide 30-12.

KM DNA-polymerase; DNAP; 5' nuclease; Thermus; DNA cleavage;

KM RNA cleavage; transcleavage; pilot oligonucleotide; hairpin; ss.

OS Synthetic.

PN W09429482-A.

PD 22-DEC-1994.

PF 06-JUN-1994; U06253.

PR 04-JUN-1993; US-073384.

PA (THIR-) THIRD WAVE TECHNOLOGIES INC.

PI Brow MAD, Dahlberg JE, Lyamchev VI;

DR WPI; 95-036504/05.

PT 5' Nuclease(s) derived from thermostable DNA polymerase(s) - have

PT cleavage activity with reduced synthetic ability, used for

PT detection of specific target sequences.

PS Disclosure; Page 98; 159pp; English.

CC The ability of a 5' nuclease to cleave a specific sequence

CC structure was tested using the pilot oligonucleotides 19-12

CC (given in 080758) and 30-12 (080758), which hybridize to

CC substrate DNA to create a short hairpin with a stabilizing

CC tetra-loop, and pilot oligonucleotide 30-0 (080760), which

CC hybridizes to substrate RNA.

SQ Sequence 42 BP; 8 A; 12 C; 12 G; 10 T;

Mar 25 02:55

US-08-644-289-5.rmg

11

Query Match 100.0%; Score 10; DB 13; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.99e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 aggcatacct 32
|||||||
Oy 1 AGGCATGCCT 10

RESULT 14

ID Q80759 standard; DNA; 42 BP.
AC Q80759;
DT 19-JUL-1995 (first entry)
DE Pilot oligonucleotide 30-12.
KM DNA-polymerase; DNP; 5' nuclease; Thermus; DNA cleavage;
KM RNA cleavage; transcleavage; pilot oligonucleotide; hairpin; ss.
OS Synthetic.
PN W09429482-A.
PD 22-DEC-1994.
PF 06-JUN-1994; U06253.
PR 04-JUN-1993; US-073384.
PI (THR-) THIRD WAVE TECHNOLOGIES INC.
PI Brow MAD, Dahlberg JE, Lyanichev VI;
DR WPI; 95-03504/05.
PT 5' Nuclease(s) derived from thermostable DNA polymerase(s) - have
PT cleavage activity with reduced synthetic ability, used for
PT detection of specific target sequences.
PS Disclosure; Page 98; 159pp; English.
CC The ability of a 5' nuclease to cleave a specific sequence
CC structure was tested using the pilot oligonucleotides 19-12
CC (given in Q80758) and 30-12 (Q80758), which hybridize to
CC substrate DNA to create a short hairpin with a stabilizing
CC tetra-loop, and pilot oligonucleotide 30-0 (Q80760), which
CC hybridizes to substrate RNA.
SQ Sequence 42 BP; 8 A; 12 C; 12 G; 10 T;

Query Match 100.0%; Score 10; DB 13; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.99e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 aggcatacct 32
|||||||
Cp 10 AGGCATGCCT 1

RESULT 15
ID Q31958 standard; DNA; 46 BP.
AC Q31958;
DT 27-APR-1993 (first entry)
DE p53 binding site of cloned human genomic fragment #211.
KM P53; DNA-binding; cancer; neoplasia; tumour; concatenate; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT misc_feature 16..25
FT /*tag= a
FT /*note= "homology with consensus p53 binding site"
FT misc_feature 26..31
FT /*tag= b
FT /*note= "homology with consensus p53 binding site"
PN EP-518650-A.
PD 16-DEC-1992.
PF 10-JUN-1992; 305333.
PR 14-JUN-1991; US-715182.
PR 31-MAR-1992; US-860758.

Mar 25 02:55

US-08-644-289-5.rmg

12

PA (PHAR-) PHARMAGENICS INC.
PA (UYJO) UNITV JOHNS HOPKINS.
PI Kinzler KM, Sherman MI, Vogelstein B;
DR WPI; 92-417505/51.
PT Detection and expression of wild type p53 protein - useful for
PT diagnosing and treating cancers, and for screening potential
PT chemotherapeutic agents

PS Example 10; Fig 10; 51pp; English.
CC Wild-type p53 protein binds specific fragments of human chromosomal
CC DNA. 23 clones which bound p53 were obt. by whole genome PCR.
CC The sequences of these 23 clones was compared and an average insert
CC of 307 bp (range 139-470) was found. Only 18 of the clones contained
CC unique fragments. Alignment of the inserts revealed a striking and
CC consistent feature, each binding site contained two copies of the
CC 10 bp motif 5'-RRRCWGGYY-3' separated by 0-13 bp. The 10 bp
CC consensus sequence contains internal symmetry. Mutation analysis
CC revealed that the dimer of the consensus motif is required to bind
CC p53 and that mutant p53 proteins found in tumours do not bind to the
CC consensus sequence. See also Q31948-84.
SQ Sequence 46 BP; 6 A; 16 C; 6 G; 18 T;

Query Match 100.0%; Score 10; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.99e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 16 aggcatacct 25
|||||||
Oy 1 AGGCATGCCT 10

Search completed: Tue Mar 25 02:57:35 1997
Job time : 14 secs.

Mar 25 02:57

US-08-644-289-5.stl

1

US-08-644-289-5 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligentics, Inc.

MPsrch_m n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Mar 25 02:57:55 1997; MsrPar time 54.05 Seconds

Tabular output not generated. 81.591 Million cell updates/sec

Title: >US-08-644-289-5
Description: (1-10) from US08644289.seq
N.A. Sequence: 1 AGCGATCCCT 10
Comp: TCCGTACGCA

Scoring table: TABLE default
Gap 10

Mismatch STD : Dbase 0; Query 0

Searched: 630489 seqs, 220513910 bases x 2

Post-Processing: Minimum Match 0%

Listing first 45 summaries

Database: EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99

Database: EST-STS-TWO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129

Mar 25 02:57

US-08-644-289-5.stl

2

130:STS2 131:STS3 132:STS4 133:STS5 134:STS6 135:STS7
136:STS8 137:STS9 138:STS10 139:STS11 140:STS12
141:STS13 142:STS14 143:STS15 144:STS16 145:STS17
146:STS18 147:STS19 148:STS20 149:STS21 150:STS22
151:STS23 152:STS24 153:STS25 154:STS26 155:STS27
156:STS28 157:STS29 158:STS30 159:STS31 160:STS32
161:STS33 162:STS34 163:STS35 164:STS36 165:STS37
166:STS38 167:STS39 168:STS40 169:STS41 170:STS42
171:STS43 172:STS44 173:STS45 174:STS46

Statistics: Mean 5.340; Variance 0.849; scale 6.287

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	DB Length	ID	Description	Pred. No.
1	10	100.0	205 158	HS128324	zab9n06.r1 Soares fet	2.46e+01
2	10	100.0	239 46	HUM052C06B	Human fetal brain cDN	2.46e+01
3	10	100.0	239 46	HUM052C06B	Human fetal brain cDN	2.46e+01
4	10	100.0	255 150	W14656	mb34g05.r1 Soares mou	2.46e+01
5	10	100.0	269 133	G11844	human STS WI-10852.	2.46e+01
6	10	100.0	272 174	HS2444WH5	H. sapiens (D1S2675) D	2.46e+01
7	10	100.0	272 174	HS2444WH5	H. sapiens (D1S2675) D	2.46e+01
8	10	100.0	290 26	H65997	yr73c10.r1 Homo sapie	2.46e+01
9	10	100.0	292 39	HSC08081	H. sapiens partial cD	2.46e+01
10	10	100.0	295 174	HS2444WH5	H. sapiens (D1S2675) D	2.46e+01
11	10	100.0	295 136	HS2444WH5	H. sapiens (D1S2675) D	2.46e+01
12	10	100.0	309 60	N46803	yy52e05.r1 Homo sapie	2.46e+01
13	10	100.0	311 92	R71722	yy52e05.r1 Homo sapie	2.46e+01
14	10	100.0	313 52	MUSG501048	Mouse 3'-directed cD	2.46e+01
15	10	100.0	313 109	T78441	EST44578 Homo sapiens	2.46e+01
16	10	100.0	319 122	T77928	yd16105.r1 Homo sapie	2.46e+01
17	10	100.0	330 174	HS2444WH5	H. sapiens (D1S2675) D	2.46e+01
18	10	100.0	330 174	HS2444WH5	H. sapiens (D1S2675) D	2.46e+01
19	10	100.0	330 166	HS2444WH5	H. sapiens (D1S2675) D	2.46e+01
20	10	100.0	331 6	DY181C11	H. sapiens partial cD	2.46e+01
21	10	100.0	340 18	H40613	Human fetal brain cDN	2.46e+01
22	10	100.0	342 46	HUM064H04A	Human fetal brain cDN	2.46e+01
23	10	100.0	359 104	T06552	EST04441 Homo sapiens	2.46e+01
24	10	100.0	365 149	W10375	ma38c02.r1 Soares mou	2.46e+01
25	10	100.0	365 149	W10375	ma38c02.r1 Soares mou	2.46e+01
26	10	100.0	380 49	HUM051B03B	Human placenta cDNA 5	2.46e+01
27	10	100.0	387 121	T75586	10464 Arabidopsis th	2.46e+01
28	10	100.0	396 136	HUM451048	Human chromosome 4 (c	2.46e+01
29	10	100.0	399 119	T67119	Human chromosome 4 (c	2.46e+01
30	10	100.0	401 153	W21554	zb52g12.r1 Soares fet	2.46e+01
31	10	100.0	412 46	HUM047E12B	Human fetal brain cDN	2.46e+01
32	10	100.0	416 46	HUM051B03B	Human placenta cDNA 5	2.46e+01
33	10	100.0	422 103	R1C54377A	Rice cDNA, partial se	2.46e+01
34	10	100.0	430 157	AT60221	21777 Arabidopsis th	2.46e+01
35	10	100.0	432 23	H55900	yr07b07.s1 Homo sapie	2.46e+01
36	10	100.0	435 127	T95171	ye39107.r1 Homo sapie	2.46e+01
37	10	100.0	440 36	H98640	yx12e06.s1 Homo sapie	2.46e+01
38	10	100.0	447 46	HUM050C08B	Human fetal brain cDN	2.46e+01
39	10	100.0	466 163	HS590321	zb69c09.s1 Homo sapie	2.46e+01
40	10	100.0	486 132	G10653	human STS CHIC UTR 03	2.46e+01
41	10	100.0	498 118	T61429	yc06a02.s1 Homo sapie	2.46e+01
42	10	100.0	502 26	H64117	yr57g04.r1 Homo sapie	2.46e+01

Mar 25 02:57

US-08-644-289-5.rst

5

REFERENCE 1 (bases 1 to 239)
AUTHORS Fujiwara, T.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-1995) to the DDBJ/EMBL/GenBank databases. Tautomu Fujiwara, Otsuka GEN Research Institute, Otsuka Pharmaceutical Co., Ltd, 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01, Japan

REFERENCE 2 (bases 1 to 239)
AUTHORS Fujiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Takachi, A., Takeda, S., Matanabe, T., Takahashi, E., Hirai, Y., Maekawa, H., Shin, S. and Nakamura, Y.
TITLE Large-scale sequencing project at Otsuka GEN Research Institute
JOURNAL Unpublished (1995)
COMMENT Submitted (7-Nov-1995) to DDBJ by:
Tautomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical CO., Ltd
463-10 Kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone: 0886-65-2888
Fax: 0886-37-1035.

FEATURES
source
1..239
Location/Qualifiers
/clone_lib="human fetal brain"
/dev_stage="Fetus"
/organism="Homo sapiens"
/sequenced_mol="cDNA to mRNA"
/issue_type="brain"

BASE COUNT 65 a 58 c 55 g 44 t 17 others
ORIGIN

Query Match 100.0%; Score 10; DB 46; Length 239;
Best Local Similarity 100.0%; Pred. No. 2.46e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 129 aggcattgct 138
|||||
Cp 10 AGGCATGCT 1

RESULT 4
LOCUS W14696 255 bp mRNA EST 26-APR-1996
DEFINITION mb34g05.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 331352
5'
ACCESSION W14696
NID 91288727
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
1 (bases 1 to 255)
REFERENCE 1 (bases 1 to 255)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT

Mar 25 02:57

US-08-644-289-5.rst

6

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:212752
Seq primer: mob.REGA+ET
High quality sequence stop: 241.

FEATURES
source
NCBI gi: 1288727
Location/Qualifiers
1..255
/organism="Mus musculus"
/note="Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TTTTACCACTGTAAGTGGACGCCGCAATTTTCTTTTCTTTT 3'), double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 49 a 69 c 85 g 52 t
ORIGIN

Query Match 100.0%; Score 10; DB 150; Length 255;
Best Local Similarity 100.0%; Pred. No. 2.46e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 20 aggcattgct 29
|||||
QY 1 AGGCATGCT 10

RESULT 5
LOCUS G11844 269 bp DNA STS 19-OCT-1995
DEFINITION human STS WI-10852.
ACCESSION G11844
NID 91022599
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human STS derived from random genomic DNA.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 269)
REFERENCE 1 (bases 1 to 269)
AUTHORS Hudson, T.
TITLE Mapped STS
JOURNAL Whitehead Institute/MIT Center for Genome Research; Physically Mapped STS
COMMENT Unpublished (1995)
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research

Mar 25 02:57

US-08-644-289-5.rst

7

Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA

Tel: 617 252 1900

Fax: 617 252 1902

Email: thudson@genome.wi.mit.edu

Primer A: TTCTTCGATCTAGGCTAAGG

Primer B: TCATTCCTACTGCGAATTTCTAGG

STS size: 200

PCR Profile:

Presoak:

Denaturation:

Annealing: 56 degrees C

Polymerization:

PCR cycles: 35

Thermal Cycler:

Protocol:

Template: 10 ng

Primer: each 5 pM

dNTPs: each 4 mM

Taq Polymerase: 0.025 units/ul

Total Vol: 20 ul

Buffer:

MgCl2: 1.5 mM

KCl: 50 mM

Tris-HCl: 10 mM

pH: 9.3

Prepared with primer pairs derived from random genomic sequence.

FEATURES

source

1..269

/location/Qualifiers

/organism="Homo sapiens"

/note="human"

STS

70..269

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

70..91

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

complement(246..269)

/map="746_D_8; 854_E_5; 891_G_9; 950_G_3; 950_G_4; 747_D_11; 811_D_12; 848_C_10; 854_E_10; 870_G_10; 730.7 cR from top of Chr1 linkage group"

Mar 25 02:57

US-08-644-289-5.rst

8

DE AFMA244wh5; single read.

KM CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;

OS Homo sapiens (human)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.

RM [1]

RP 1-272

RA Weissenbach J.;

RT ;

RL Submitted (01-SEP-1995) to the EMBL/GenBank/DBJ databases.

RL Genethon, B.P. 60, 91002 Evry Cedex France. E-mail:

RL Jean.Weissenbach@genethon.fr

RM [2]

RP 1-272

RA Dib C., Faure S., Fizames C., Samson D., Drouot N., Vignal A.,

RA Millaeseau P., Marc S., Hazan J., Seboun E., Lathrop M.,

RA Gyapay G., Morissette J., Weissenbach J.;

RT "A comprehensive genetic map of the human genome based on 5,264

microsatellites";

RL Nature 380:152-154 (1996).

CC full automatic;

FH Key Location/Qualifiers

FH

FT source

FT 1..272

/organism="Homo sapiens"

/cell_line="CEPH 134702"

/clone_lib="genomic DNA"

/chromosome="1"

FT /note="cloning vector is M13mp18"

SQ Sequence 272 BP; 82 A; 79 G; 64 G; 44 T; 3 other;

Query Match

Best Local Similarity 100.0%; Score 10; DB 174; Length 272;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 141 agcgcgcct 150

|||||

1 AGCGATGCCT 10

Qy

RESULT

ID HSA244wh5

standard; DNA; STS; 272 BP.

AC 252679;

DT 18-MAR-1996 (Rel. 47, Created)

DT 23-MAR-1996 (Rel. 47, Last updated, Version 3)

DE H.sapiens (DIS2675) DNA segment containing (CA) repeat; clone

DE AFMA244wh5; single read.

DE CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;

KM microsatellite marker; repeat polymorphism, STS.

OS Homo sapiens (human)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.

RM [1]

RP 1-272

RA Weissenbach J.;

RT ;

RL Submitted (01-SEP-1995) to the EMBL/GenBank/DBJ databases.

RL Genethon, B.P. 60, 91002 Evry Cedex France. E-mail:

RL Jean.Weissenbach@genethon.fr

RM [2]

RP 1-272

RA Dib C., Faure S., Fizames C., Samson D., Drouot N., Vignal A.,

RA Millaeseau P., Marc S., Hazan J., Seboun E., Lathrop M.,

Mar 25 02:57

US-08-644-289-5.1st

9

RA Gyapay G., Morissette J., Weisenbach J.;
RT "A comprehensive genetic map of the human genome based on 5,264
microsatellites";
RL Nature 380:152-154(1996).
CC full automatic;
FH Key Location/Qualifiers
FT source 1..272
FT /organism="Homo sapiens"
FT /cell_line="CEPH 134702"
FT /clone_lib="genomic DNA"
FT /chromosome="1"
FT /note="cloning vector is M13mp18"
SQ Sequence 272 BP; 82 A; 79 C; 64 G; 44 T; 3 other;

Query Match 100.0%; Score 10; DB 174; Length 272;
Best Local Similarity 100.0%; Pred. No. 2.46e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 141 aggcatacct 150
|||||
Cp 10 AGGCATGCCT 1

RESULT 8
LOCUS H65997 290 bp mRNA EST 18-OCT-1995
DEFINITION vt73c10.r1 Homo sapiens cDNA clone 210930 5'.
ACCESSION H65997
MID 91024737
KEYWORDS EST.
SOURCE human clone=210930 primer=M13R1 library=Soares fetal liver spleen
INFUS vector=PT773D (Pharmacia) with a modified polyLinker
host=DH10B (ampicillin resistant) Rsite1=Pac I Rsite2=Eco RI Liver
and spleen from a 20 week-post conception male fetus. 1st strand
cDNA was primed with a Pac I - oligo(dT) primer [5'
AACTGCAATATTAATTAAGATCTTTTCTTTTCTTTT 3'], double-stranded
cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
I and cloned into the Pac I and Eco RI sites of the modified PT773
vector. Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM

Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 290)

REFERENCE

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Maier, M.,
Parsons, V., Rifkin, B., Rohlfing, T., Soares, M., Tan, F.,
Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
TITLE The Mashu-Merck EST Project
JOURNAL Unpublished (1995)

COMMENT
Contact: Wilson RK
Mashu-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 189
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the

Mar 25 02:57

US-08-644-289-5.1st

10

FEATURES
source
1..290
/organism="Homo sapiens"
/clone="210930"
/note="human"
IMGE Consortium (info@image.llnl.gov) for further information.
Location/Qualifiers

BASE COUNT 77 a 64 c 71 g 73 t 5 others
ORIGIN
mRNA
<1..>290

Query Match 100.0%; Score 10; DB 26; Length 290;
Best Local Similarity 100.0%; Pred. No. 2.46e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 206 aggcatacct 215
|||||
Qy 1 AGGCATGCCT 10

RESULT 9
LOCUS HSC0E081 292 bp RNA EST 05-NOV-1994
DEFINITION H. sapiens partial cDNA sequence; clone c-0e08.
ACCESSION Z42301
MID 9565710
KEYWORDS partial cDNA sequence; transcribed sequence fragment.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Catarrhini; Hominoidea; Homo.
1 (bases 1 to 292)

REFERENCE

AUTHORS Geneexpress.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-1994) to the EMBL/GenBank/DBJ databases.
Genetion, B.P. 60, 91002 Evry Cedex France and Genetique
Moleculaire et Biologie du developpement, CNRS UPR420 B.P. 8, 94801
Villejuif Cedex France. E-mail: geneexpress@genetion.fr

REFERENCE

AUTHORS Geneexpress.
TITLE The Geneexpress cDNA program
JOURNAL Unpublished
3 (bases 1 to 292)

REFERENCE

AUTHORS Devignes, M.D., Duprat, S., Houlligat, R., Jumeau, M.N., Lamy, B.,
Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,
Sebastien-Rabackich, C. and Tessier, A.
TITLE IMAGE: Integrated molecular analysis of the human genome and its
expression
JOURNAL C.R. Acad. Sci., III, Sci. Vie 318, 263-272 (1995)
COMMENT Clone library from B.Souares, Psychiatry Dept. Columbia University
USA.

Cloning method: total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the lsmid BA
vector;
Sequencing method: single read, full automatic;
Primer: M13 reverse
cDNA sequence colinear to mRNA
Stretch_removed: nothing
Normalization method: Bento Soares, P.N.A.S in press;
Geneexpress_library_idt: C;
Geneexpress_sequence_idt: Y1c-0e08;

Mar 25 02:57

US-08-644-289-5.rst

11

No significant homology found with :
genbank release 81 swissprot release 28.
Location/Qualifiers

FEATURES
source

1..292

/organism="Homo sapiens"
/dev stage="3 months old"
/isolate="muscular atrophy patient"
/tissue type="total brain"
/clone_lib="normalized infant brain cDNA"
/sex="Female"

BASE COUNT 79 a 63 c 53 g 96 t 1 others

ORIGIN

Query Match 100.0%; Score 10; DB 39; Length 292;
Best Local Similarity 100.0%; Pred. No. 2.46e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 266 agcagtcct 275

10 AGCATGCTT 1

RESULT 10

HS3122C9 standard; DNA; STS; 295 BP.

AC 252895;

DT 18-MAR-1996 (Rel. 47, Created)

DT 23-MAR-1996 (Rel. 47, last updated, Version 3)

DE H.sapiens (D17S1818) DNA segment containing (CA) repeat; clone

DE AFMA312c9; single read.

KM CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;

KM microsatellite marker; repeat polymorphism; STS.

OC Homo sapiens (human)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.

RP 1-295

RA Weissenbach J.;

RT Submitted (01-SEP-1995) to the EMBL/GenBank/DBJ databases.

RL Genethon, B.P. 60, 91002 Evry Cedex France. E-mail:

RL Jean.Weissenbach@genethon.fr

RN 121

RP 1-295

RA D1b C., Faure S., Fizames C., Samson D., Drouot N., Vignal A.,

RA Millaesau P., Marc S., Hazan J., Seboun E., Lathrop M.,

RA Gyapay G., Morissette J., Weissenbach J.;

RT "A comprehensive genetic map of the human genome based on 5,264

RT microsatellites";

CC Nature 380:152-154(1996).

CC full automatic;

CC Key

FT source

FT 1..295

FT /organism="Homo sapiens"

FT /cell_line="CEPH 134702"

FT /clone_lib="genomic DNA"

FT /chromosome="17"

FT /note="Cloning vector is M13mp18"

Query Match 100.0%; Score 10; DB 174; Length 295;

Best Local Similarity 100.0%; Pred. No. 2.46e+01;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Mar 25 02:57

US-08-644-289-5.rst

12

Db 151 agcagtcct 160
10 AGCATGCTT 1

RESULT 11

LOCUS HS3122C9 295 bp DNA STS 23-MAR-1996

DEFINITION H.sapiens (D17S1818) DNA segment containing (CA) repeat; clone

AFMA312c9; single read.

ACCESSION 252895

KEYWORDS 91234195

CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;

microsatellite marker; repeat polymorphism; STS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 295)

Submitted (01-SEP-1995) Genethon, B.P. 60, 91002 Evry Cedex France.

E-mail: Jean.Weissenbach@genethon.fr

AUTHORS

D1b C., Faure S., Fizames C., Samson D., Drouot N., Vignal A.,

Millaesau P., Marc S., Hazan J., Seboun E., Lathrop M., Gyapay G.,

Morissette J. and Weissenbach J.;

A comprehensive genetic map of the human genome based on 5,264

microsatellites

Nature 380, 152-154 (1996)

full automatic.

FEATURES

source

1..295

/organism="Homo sapiens"

/note="Cloning vector is M13mp18"

/cell_line="CEPH 134702"

/clone_lib="genomic DNA"

/chromosome="17"

BASE COUNT 65 a 98 c 45 g 81 t 6 others

ORIGIN

Query Match 100.0%; Score 10; DB 136; Length 295;
Best Local Similarity 100.0%; Pred. No. 2.46e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 151 agcagtcct 160

10 AGCATGCTT 1

RESULT 12

LOCUS N46803 309 bp mRNA EST 14-FEB-1996

DEFINITION Y52605.r1 Homo sapiens cDNA clone 277184 5'.

ACCESSION N46803

KEYWORDS 91187969

human clone=277184 primer=T7 library=Soares multiple sclerosis

2NBHSP vector=PTT30 (Pharmacia) with a modified polylinker

V type: phagemid host=DH10B (ampicillin resistant) Rst1=Not I

Rst2=Eco RI 46 year old male. 1st strand cDNA was primed with a

Not I - oligo (dT) primer

15'-TCTTACCACTGCAAGTCGACCGCGCATTTTCTTTTCTTTT-3',

double-stranded cDNA was size selected, ligated to Eco RI adapters

(Pharmacia), digested with Not I and cloned into the Not I and Eco

Mar 25 02:57

US-08-644-289-5.rst

13

RI sites of a modified pT7 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH).

ORGANISM

Homo sapiens

REFERENCE

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 309)

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE

The WashU-Merck EST Project
Unpublished (1995)

JOURNAL

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 259
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

COMMENT

Location/Qualifiers
1..309
/organism="Homo sapiens"
/clone="277184"
/note="human"
<1..>309

FEATURES

source

1..309
/organism="Homo sapiens"
/clone="277184"
/note="human"
<1..>309

BASE COUNT

75 a 66 c 67 g 98 t 3 others

ORIGIN

Query Match 100.0%; Score 10; DB 60; Length 309;
Best Local Similarity 100.0%; Pred. No. 2.46e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 139 agcagtcgct 148
|||||
Cp 10 AGCATGCCT 1

RESULT 13

LOCUS R71722 311 bp mRNA EST 02-JUN-1995
DEFINITION yj85g05.c1 Homo sapiens cDNA clone 155576 5'.
ACCESSION R71722
NID 9845754
KEYWORDS EST.
SOURCE human clone=155576 library=Soares breast 2NbhBat vector=pT7T3D (Pharmacia) with a modified polyLinker host=DH10B (ampicillin resistant) primer=M13RP1 Reitel2Not I Reitel2-Eco RI Adult female. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGACGCCGCCCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 230. Library constructed by Bento Soares and M.Fatima Bonaldo.

Mar 25 02:57

US-08-644-289-5.rst

14

ORGANISM

Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 311)

REFERENCE

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

AUTHORS

The WashU-Merck EST Project
Unpublished (1995)

TITLE

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 189
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

JOURNAL

Location/Qualifiers
1..311
/organism="Homo sapiens"
/clone="155576"
/note="human"

COMMENT

Query Match 100.0%; Score 10; DB 92; Length 311;
Best Local Similarity 100.0%; Pred. No. 2.46e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FEATURES

source

1..311
/organism="Homo sapiens"
/clone="155576"
/note="human"

BASE COUNT

69 a 101 c 63 g 74 t 4 others

ORIGIN

Query Match 100.0%; Score 10; DB 92; Length 311;
Best Local Similarity 100.0%; Pred. No. 2.46e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 220 agcagtcgct 229
|||||
Qy 1 AGCATGCCT 10

RESULT 14

LOCUS MUSGS01048 313 bp mRNA EST 09-DEC-1995
DEFINITION Mouse 3'-directed cDNA, MUSGS01048, clone mc0750.
ACCESSION D18866
NID g1089499
KEYWORDS EST(expressed sequence tag); Gene signature(GS); development; transcribed sequence.
SOURCE Mus musculus (strain C57BL/6Jsub species domesticus,) decidua tissue (day 6.5-8.5 of gestation) cDNA to mRNA.
ORGANISM Mus musculus

REFERENCE

Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathu; Myomorpha; Muridae; Murinae; Mus.
1 (bases 1 to 313)
Kawamoto, S., Okubo, K., Yoshii, J., Katsuki, M. and Matsubara, K. Analysis of gene expression in mouse embryogenesis by 3'-directed cDNA sequencing
Unpublished (1993)
Submitted (10-Sep-1993) to DDBJ by:
Shoko Kawamoto
Institute for Cellular and Molecular Biology
Osaka University

AUTHORS

Unpublished (1993)

TITLE

Submitted (10-Sep-1993) to DDBJ by:
Shoko Kawamoto
Institute for Cellular and Molecular Biology
Osaka University

JOURNAL

Submitted (10-Sep-1993) to DDBJ by:
Shoko Kawamoto
Institute for Cellular and Molecular Biology
Osaka University

COMMENT

Submitted (10-Sep-1993) to DDBJ by:
Shoko Kawamoto
Institute for Cellular and Molecular Biology
Osaka University

Submitted (10-Sep-1993) to DDBJ by:
Shoko Kawamoto
Institute for Cellular and Molecular Biology
Osaka University

Mar 25 02:57

US-08-644-280-5.txt

15

3-1, Yamadaoka
Saita, Osaka, 565

Japan

Phone: 06-879-7992

Fax: 06-877-1922

Email: shokohimb.osaka-u.ac.jp.

FEATURES

source

1..313

/organism="Mus musculus"

/strain="C57BL/6J"

/sub_species="domesticus"

/sequenced_mol="cDNA to mRNA"

/tissue_type="decidual tissue (day 6.5-8.5 of gestation)"

BASE COUNT

74 a

86 c

91 g

62 t

ORIGIN

Query Match

100.0%; Score 10; DB 52; Length 313;

Best Local Similarity 100.0%; Pred. No. 2,46e+01;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 184 aggcattcct 193

Cp 10 AGGCATGCCT 1

RESULT 15

LOCUS T28441 313 bp mRNA EST 06-SEP-1995

DEFINITION EST44578 Homo sapiens cDNA 5' end similar to

N-acetylglucosaminyl(beta 1-6) transferase 1 (HT:992).

ACCESSION T28441

NID 9610539

KEYWORDS EST.

SOURCE human primer=M13 Reverse library=Human Brain.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

1 (bases 1 to 313)

Mar 25 02:57

US-08-644-289-5.txt

16

Fax: 3018699423

Email: tdbinfo@tdb.tigr.org

For clone availability, additional sequence and expression

information related to this EST, please contact the TIGR Database

(tdbinfo@tdb.tigr.org).

FEATURES

source

1..313

/organism="Homo sapiens"

/note="human"

BASE COUNT 84 a 64 c 65 g 100 t

ORIGIN

Query Match 100.0%; Score 10; DB 109; Length 313;

Best Local Similarity 100.0%; Pred. No. 2,46e+01;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 126 aggcattcct 135

Cp 10 AGGCATGCCT 1

Search completed: Tue Mar 25 02:58:17 1997
Job time : 62 secs.